

Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 143
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-8

Query Match 100.0%; Score 554; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.7e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 48 KTCYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 3
US-10-106-698-6266
Sequence 6266, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005F1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 6266
LENGTH: 337
TYPE: PRT
ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match 100.0%; Score 554; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 4.5e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 74 KTCYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 133
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 134 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 169

RESULT 4
US-09-880-503-3
Sequence 3, Application US/09880503
Patent No. US20020131964A1

GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 100.0%; Score 554; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.6e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 48 KTCYEGNGHYFGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 96
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPEE 143

RESULT 5
US-09-264-468B-1
Sequence 1, Application US/09264468B
Patent No. US20020106775A1
GENERAL INFORMATION:
APPLICANT: Wang, Jieyi
APPLICANT: Nienaber, Vicki L.
APPLICANT: Henkin, Jack
APPLICANT: Smith, Richard A.
APPLICANT: Walter, Karl A.
APPLICANT: Severin, Jean M.
APPLICANT: Edalji, Rohinton
APPLICANT: Johnson Jr., Robert W.
APPLICANT: Holzman, Thomas F.
TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
FILE REFERENCE: 6310.US.P1
CURRENT APPLICATION NUMBER: US/09/264,468B
CURRENT FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: US 09/036,361
PRIOR FILING DATE: 1998-03-06
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(20)
OTHER INFORMATION: Leader sequence
NAME/KEY: VARIANT
LOCATION: (279)...(279)
OTHER INFORMATION: Xaa = any amino acid
NAME/KEY: VARIANT
LOCATION: (302)...(302)
OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 100.0%; Score 554; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

```

RESULT 6
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: WPM01-029P3ENM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-301-822-161

```

Query Match	100.0%;	Score 554;	DB 12;	Length 431;
Best Local Similarity	100.0%;	Pred. No. 5.9e-56;		
Matches 96;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	KTCYEGNGHYRGKASDTTGRPCLPWNSATVLTQTYHAHRSALQLGKKNYCRP	60	
Db	68	KTCYEGNGHYRGKASDTTGRPCLPWNSATVLTQTYHAHRSALQLGKKNYCRP	127	
Qy	61	RRRPWCYVQGLKPLVQECMVHDCADGKXPSSP	96	
Db	128	RRRPWCYVQGLKPLVQECMVHDCADGKXPSSP	163	

```

RESULT 7
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication NO. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Ocleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; PRIORITY FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8

```

```

; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: prt
; ORGANISM: Homo sapiens
US-10-131-985-21

```

	Query Match	100.0%	Score 554;	DB 12;	Length 431;
	Best Local Similarity	100.0%;	Prod No. 5.9e-56;		
	Matches 96;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	KTCYEGNGHFRGKASTD	TMGRCPLPNSATVTLQOYTHAHRSDALQLGLGKHNYCRNP	60	
Db	68	KTCYEGNGHFRGKASTD	TMGRCPLPNSATVTLQOYTHAHRSDALQLGLGKHNYCRNP	127	
Qy	61	RRRPWCYYQVGLKPLVQECMVHDCADGKTPSSPPEE	96		
Db	128	RRRPWCYYQVGLKPLVQECMVHDCADGKTPSSPPEE	163		

```

RESULT 8
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication NO. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076.421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

```

	Query Match	100.0%	Score 554;	DB 14;	Length 431;
	Best Local Similarity	100.0%;	Pred. NO. 5.9e-56;		
	Matches 96;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	KTCYEGNGHFYRGKASTD	TGRCPLPNSATVLOQTY	HAHRSDALQLGLGKHNYCR	PDN 60
Db	68	KTCYEGNGHFYRGKASTD	TGRCPLPNSATVLOQTY	HAHRSDALQLGLGKHNYCR	PDN 127
Qy	61	RRRPWCYQVGLKPLVQEC	MDADGKKPSSPPEE	96	
Db	128	RRRPWCYQVGLKPLVQEC	MDADGKKPSSPPEE	163	

```

RESULT 9
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

```



```
/ TITLE OF INVENTION: OF CERVICAL CANCER
/ FILE REFERENCE: MRI-035
/ CURRENT APPLICATION NUMBER: US/10/171,311
/ CURRENT FILING DATE: 2002-06-12
/ PRIOR APPLICATION NUMBER: US 60/298,159
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/298,155
/ PRIOR FILING DATE: 2001-06-13
/ PRIOR APPLICATION NUMBER: US 60/335,936
/ PRIOR FILING DATE: 2001-11-14
/ NUMBER OF SEQ ID NOS: 238
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 184
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-171-311-184

Query Match 100.0%; Score 554; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 163

RESULT 10
US-10-193-656-4
/ Sequence 4, Application US/10193656
/ Publication No. US20030096733A1
/ GENERAL INFORMATION:
/ APPLICANT: NY, Tor
/ APPLICANT: HOLMDAHL, Rikard
/ APPLICANT: LI, Jinan
/ TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
/ FILE REFERENCE: 3810/1J577-US3
/ CURRENT APPLICATION NUMBER: US/10/193,656
/ CURRENT FILING DATE: 2002-07-10
/ PRIOR APPLICATION NUMBER: US 60/304,461
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/304,490
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: US 60/305,182
/ PRIOR FILING DATE: 2001-07-13
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: GenBank / P00749
/ DATABASE ENTRY DATE: 1986-07-21
/ RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match 100.0%; Score 554; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 163
```

```
RESULT 11
US-10-247-671-149
/ Sequence 149, Application US/10247671
/ Publication No. US20030194721A1
/ GENERAL INFORMATION:
/ APPLICANT: Mikita, Thomas
/ APPLICANT: Shiffman, Dov
/ APPLICANT: Porter, Gordon, J.
/ APPLICANT: Kaser, Matthew R.
/ TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
/ FILE REFERENCE: PA-0050 US
/ CURRENT APPLICATION NUMBER: US/10/247,671
/ CURRENT FILING DATE: 2002-09-18
/ PRIOR APPLICATION NUMBER: 60/323,784
/ PRIOR FILING DATE: 2001-09-19
/ NUMBER OF SEQ ID NOS: 186
/ SOFTWARE: PERL Program
/ SEQ ID NO 149
/ LENGTH: 431
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match 99.6%; Score 552; DB 12; Length 431;
Best Local Similarity 99.0%; Pred. No. 1e-55;
Matches 95; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 96
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 163

RESULT 12
US-10-407-821-2
/ Sequence 2, Application US/10407821
/ Publication No. US20030219386A1
/ GENERAL INFORMATION:
/ APPLICANT: IDELL, STEVEN
/ TITLE OF INVENTION: INTRAEPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
/ TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
/ TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
/ FILE REFERENCE: UTSN:022US
/ CURRENT APPLICATION NUMBER: US/10/407,821
/ CURRENT FILING DATE: 2003-04-04
/ PRIOR APPLICATION NUMBER: 60/414,202
/ PRIOR FILING DATE: 2002-09-27
/ PRIOR APPLICATION NUMBER: 60/370,466
/ PRIOR FILING DATE: 2002-04-05
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 411
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-407-821-2

Query Match 98.2%; Score 544; DB 12; Length 411;
Best Local Similarity 99.0%; Pred. No. 8.1e-55;
Matches 95; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPPPE 96
```

```
Db 108 RRRPCYVQVGLKLLVQECMVHDCADGKPKSSPPEE 143
|||||
RESULT 13
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1
Query Match 92.1%; Score 510; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.3e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
Db 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88
Db 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88

RESULT 14
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4
Query Match 92.1%; Score 510; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPCYVQVGLKPLVQECMVHDCADGK 135

Db 108 RRRPCYVQVGLKLLVQECMVHDCADGKPKSSPPEE 143
|||||
RESULT 15
US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittou, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, COMPOSITION
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12
Query Match 92.1%; Score 510; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOITYHAHRSDALQLGLGKHNYCRNPDN 110
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPCYVQVGLKPLVQECMVHDCADGK 138

RESULT 16
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
```

GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12
Query Match 92.1%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 110
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRPWCYVQVGLKPLVQECMVHDCADGK 138
RESULT 17
US-10-237-708-12
Sequence 12, Application US/10237708
Publication No. US20030036170A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard

Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,708
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12
Query Match 92.1%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 110
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRPWCYVQVGLKPLVQECMVHDCADGK 138
RESULT 18
US-10-237-866-12
Sequence 12, Application US/10237866
Publication No. US20030036171A1
GENERAL INFORMATION:
APPLICANT: Fleer, Reinhard
Fournier, Alain
Guitton, Jean-Dominique
Jung, Gerard
Yeh, Patrice
TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
CONTAINING SAID POLYPEPTIDES
NUMBER OF SEQUENCES: 36


```
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12
Query Match 92.1%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 51 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 110

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 21
US-09-880-503-6
Sequence 6, Application US/09880503
Patent No. US20020131964A1
GENERAL INFORMATION:
APPLICANT: CINES, Douglas B
APPLICANT: HIGAZI, Abd Al-Roof
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
TITLE OF INVENTION: TISSUE CONTRACTABILITY
FILE REFERENCE: 9596-331
CURRENT APPLICATION NUMBER: US/09/880,503
CURRENT FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/212,847
PRIOR FILING DATE: 2000-06-20
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-09-880-503-6
Query Match 92.1%; Score 510; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 7e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 48 KTCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 107

us-09-880-503-9.rapb
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 22
US-09-987-457-18
Sequence 18, Application US/09987457
Publication No. US20030013150A1
GENERAL INFORMATION:
APPLICANT: Manosroi, Aranya
APPLICANT: Manosroi, Jiradej
APPLICANT: Tayapiwatana, Chatchai
APPLICANT: Goetz, Friedrich
APPLICANT: Werner, Rolf-Guenther
TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
FILE REFERENCE: 0652.2180001
CURRENT APPLICATION NUMBER: US/09/987,457
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/268,573
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: GB 00 27 782.2
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18
LENGTH: 527
TYPE: PRT
ORGANISM: Homo sapiens (tPA)
US-09-987-457-18
Query Match 40.8%; Score 226; DB 11; Length 527;
Best Local Similarity 47.7%; Pred. No. 8.4e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 61
DB 91 TCYEGNGHYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 150

QY 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 87
DB 151 RRRPWCYVQVGLKPLVQECMVHDCADGK 176

RESULT 23
US-09-987-455-19
Sequence 19, Application US/09987455
Publication No. US20030049729A1
GENERAL INFORMATION:
APPLICANT: Aranya Manosroi
APPLICANT: Jiradej Manosroi
APPLICANT: Chatchai Tayapiwatana
APPLICANT: Friedrich Goetz
APPLICANT: Rolf-Guenther Werner
TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules
FILE REFERENCE: 0652.2190001
CURRENT APPLICATION NUMBER: US/09/987,455
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/268,574
PRIOR FILING DATE: 2001-02-15
PRIOR APPLICATION NUMBER: GB 0027779.8
PRIOR FILING DATE: 2000-11-14
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 527
TYPE: PRT
ORGANISM: Homo sapiens
US-09-987-455-19
Query Match 40.8%; Score 226; DB 11; Length 527;
```

```
Best Local Similarity 47.7%; Pred. No. 8.4e-18; Mismatches 36; Indels 0; Gaps 0;
Matches 41; Conservative 9;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHNYCRNPDR 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDOGISYRGTWSTAESGAECTNNSSALAAQKPYSGRRPDAILRGLGNHNYCRNPDR 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 91 TCYEDOGISYRGTWSTAESGAECTNNSSALAAQKPYSGRRPDAILRGLGNHNYCRNPDR 150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 151 SKPWCYVFRKAGKYSSEFCSTPACSEG 176
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 24
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only);
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 40.8%; Score 226; DB 9; Length 562;
Best Local Similarity 47.7%; Pred. No. 9e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHNYCRNPDR 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDOGISYRGTWSTAESGAECTNNSSALAAQKPYSGRRPDAILRGLGNHNYCRNPDR 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 186 SKPWCYVFRKAGKYSSEFCSTPACSEG 211
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 25
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 40.8%; Score 226; DB 10; Length 562;
Best Local Similarity 47.7%; Pred. No. 9e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;
```

```
QY 2 TCYEGNGHYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHNYCRNPDR 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDOGISYRGTWSTAESGAECTNNSSALAAQKPYSGRRPDAILRGLGNHNYCRNPDR 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 186 SKPWCYVFRKAGKYSSEFCSTPACSEG 211
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 26
US-10-443-701-4
; Sequence 4, Application US/10443701
; Publication No. US20030199016A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/10/443,701
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: US/09/703,695
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: US 60/163,607
; PRIOR FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 4
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-443-701-4

Query Match 40.8%; Score 226; DB 12; Length 562;
Best Local Similarity 47.7%; Pred. No. 9e-18;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQGLGKHNYCRNPDR 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 TCYEDOGISYRGTWSTAESGAECTNNSSALAAQKPYSGRRPDAILRGLGNHNYCRNPDR 185
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 186 SKPWCYVFRKAGKYSSEFCSTPACSEG 211
    :||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 27
US-10-193-656-8
; Sequence 8, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00750
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)...(562)
US-10-193-656-8
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 21.2586 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNELHQVPSNCDCLNGTGV.....VSHFLPWIRSHTKENGAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2301	100.0	411	1	US-08-087-163-1
2	2301	100.0	411	1	US-08-286-748B-18
3	2301	100.0	411	1	US-08-153-799-18
4	2301	100.0	431	4	US-09-101-272G-1
5	2301	100.0	431	6	5188829-1
6	2298	99.9	430	1	US-07-942-157A-3
7	2291	99.6	411	3	US-09-181-816-1
8	2285.5	99.3	430	6	5219569-2
9	2277	99.0	411	2	US-08-560-098A-48
10	2023	87.9	432	2	US-08-560-098A-47
11	2022	87.9	365	1	US-08-093-741-83
12	2022	87.9	365	1	US-08-720-012-83
13	2022	87.9	393	2	US-08-560-098A-44
14	2022	87.9	393	3	US-08-967-024C-24
15	2022	87.9	393	3	US-08-967-024C-25
16	1507.5	65.5	306	2	US-08-560-098A-45
17	1507.5	65.5	331	2	US-08-560-098A-46
18	1382	60.1	253	3	US-08-944-483-73
19	1374	59.7	254	2	US-08-560-098A-49
20	1041	45.2	200	4	US-09-101-272G-73
21	883	38.4	157	3	US-08-142-590B-25
22	871.5	37.9	527	2	US-08-811-949-39
23	864.5	37.6	527	1	US-07-609-510B-16
24	864.5	37.6	527	5	PCT-US91-01025A-2
25	864.5	37.6	527	6	5185259-8
26	864.5	37.6	562	2	US-08-811-949-43
27	864.5	37.6	562	2	US-08-560-098A-50

28	864.5	37.6	562	2	US-08-883-795A-38	Sequence 38, Appl
29	864.5	37.6	562	6	5185259-3	Patent No. 5185259
30	864.5	37.6	562	6	5200340-2	Patent No. 5200340
31	864.5	37.6	562	6	5344773-2	Patent No. 5344773
32	863.5	37.5	477	2	US-08-560-098A-51	Sequence 51, Appl
33	863.5	37.5	527	6	5520913-1	Patent No. 5520913
34	857.5	37.3	562	6	5244676-5	Patent No. 5244676
35	857	37.2	208	4	US-09-101-272G-98	Sequence 98, Appl
36	804	34.9	355	2	US-08-811-949-59	Sequence 59, Appl
37	799	34.7	437	2	US-08-811-949-51	Sequence 51, Appl
38	796	34.6	437	2	US-08-811-949-57	Sequence 57, Appl
39	793	34.5	138	2	US-08-797-689-12	Sequence 12, Appl
40	792	34.4	472	2	US-08-811-949-63	Sequence 63, Appl
41	789	34.3	437	2	US-08-811-949-55	Sequence 55, Appl
42	788	34.2	194	4	US-09-101-272G-80	Sequence 80, Appl
43	788	34.2	201	4	US-09-101-272G-96	Sequence 96, Appl
44	788	34.2	437	2	US-08-811-949-49	Sequence 49, Appl
45	787	34.2	355	2	US-08-811-949-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewich, Victor
; TITLE OF INVENTION: PRO-UKINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA: US/08/087.163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match 100.0%; Score 2301; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 2.1e-192;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYGNKGFVRG 60
|||||

Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNASATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNASATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180
Db 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTGEMKFEVENLI 240
Db 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTGEMKFEVENLI 240
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
Db 241 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 360
Db 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 360
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411
Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411
RESULT 2
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewicz
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-286-748B-18
Query Match 100.0%; Score 2301; DB 1; Length 411;

Best Local Similarity 100.0%; Pred. No. 2.le-192;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFKGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNASATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNASATVLOQTYHAHRSADALQGLGKHNYCRNPDNRPRPWCYVQVGLK 120
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180
Db 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAALYRRH 180
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTGEMKFEVENLI 240
Db 181 RGSVTVYCGGSLISPCWVISATHCFIDYPPKEDYIVYLGRSRLNSNTGEMKFEVENLI 240
QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
Db 241 LHKDYSADTLAHHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 360
Db 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKTDSCQDSSGGL 360
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411
Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHRSHTEENGLAL 411
RESULT 3
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832

TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (908) 665 2400
 ; TELEFAX: (908) 771 6159
 ; TELEX: 219484
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 411 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-153-799-18

Query Match
 Best Local Similarity 100.0%; Score 2301; DB 1; Length 411;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHCNCRNPNRRRPPWCYVQVGLK 120
 DB 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHCNCRNPNRRRPPWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
 DB 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
 QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
 DB 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
 QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
 DB 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
 QY 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPL 360
 DB 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPL 360
 QY 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411
 DB 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 4
 US-09-101-272G-1
 ; Sequence 1, Application US/09101272G
 ; Patent No. 6509445
 ; GENERAL INFORMATION:
 ; APPLICANT: Nissin Food Products Co., Ltd.
 ; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 ; FILE REFERENCE: Q50979
 ; CURRENT APPLICATION NUMBER: US/09/101,272G
 ; CURRENT FILING DATE: 1998-07-08
 ; PRIOR APPLICATION NUMBER: JP 1059/1996
 ; PRIOR FILING DATE: 1996-01-08
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: mat peptide
 ; LOCATION: (21)...()
 ; OTHER INFORMATION:
 ; NAME/KEY: misc.feature
 ; LOCATION: (20)...()
 ; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
 ; US-09-101-272G-1

Query Match
 Best Local Similarity 100.0%; Score 2301; DB 4; Length 431;

Best Local Similarity 100.0%; Pred. No. 2.2e-192;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHCNCRNPNRRRPPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHCNCRNPNRRRPPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
 DB 141 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
 QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
 DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
 DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 320
 QY 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPL 360
 DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPL 380
 RESULT 5
 5188829-1
 ; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
 ; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
 ; NUMBER OF SEQUENCES: 23
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/340,007
 ; FILING DATE: 18-AUG-1988
 ; SEQ ID NO: 1;
 ; LENGTH: 431
 5188829-1

Query Match
 Best Local Similarity 100.0%; Score 2301; DB 6; Length 431;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHCNCRNPNRRRPPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHCNCRNPNRRRPPWCYVQVGLK 140
 QY 121 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
 DB 141 PLVQECMVHDCADGKKPSSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200
 QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 240
 DB 201 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLI 260
 QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 300
 DB 261 LHKDYSADTLAHHNDIALLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCETITGPK 320
 QY 301 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPL 360
 DB 321 ENSTDYLYPEQLKMTVVKLIISHRECOQPHYGYSEVTTKMLCAADPOWKTDSCQDGGPL 380

QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYVTRVSHFLPWIRSHTKBENGIAL 411
Db 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYVTRVSHFLPWIRSHTKBENGIAL 431

RESULT 6

US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

Query Match 99.9%; Score 2298; DB 1; Length 430;
Best Local Similarity 99.8%; Pred. No. 4e-192;
Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 20 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 80 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGKKPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
Db 140 PLVQECMVHDCADGKKPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 199

QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240
Db 200 RGSVTVYVCGSLMSPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 259
QY 241 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 300
Db 260 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 319
QY 301 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGGL 360
Db 320 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGGL 379
QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYVTRVSHFLPWIRSHTKBENGIAL 411
Db 380 VCSLQGRMTLTGIVSWGRGCALKDKPGVYVTRVSHFLPWIRSHTKBENGIAL 430

RESULT 7

US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; FILE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 32904200300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 99.6%; Score 2291; DB 3; Length 411;
Best Local Similarity 99.8%; Pred. No. 1.5e-191;
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKKPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
Db 121 LLVQECMVHDCADGKKPSSPPEELKFQCGQKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240
Db 181 RGSVTVYVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSRLNSNTQGMKFEVENLI 240
QY 241 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 300
Db 241 LHKDYSADTLAHHNDIALLKIRSKEGRCQAQPSRTTQTICLPSMYNDPQGTSCETITGFGK 300
QY 301 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGGL 360
Db 301 ENSTDYLYPEOLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGGGL 360
QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYVTRVSHFLPWIRSHTKBENGIAL 411
Db 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYVTRVSHFLPWIRSHTKBENGIAL 411

RESULT 8
5219569-2

```

; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO:2
; LENGTH: 430
5219569-2

```

```

Query Match          99.38; Score 2285.5; DB 6; Length 430;
Best Local Similarity 99.88; Pred. No. 4.9e-191;
Matches 410; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIEDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWSNATVILQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSNATVILQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRPKIIGGFTTIENQFWFAAIYRRH 180
DB 141 PLVQECMVHDCADG- KPSSPPEELKFCQCGKTLRPRPKIIGGFTTIENQFWFAAIYRRH 199

QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 240
DB 200 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 259

QY 241 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFQTSCEITGFGK 300
DB 260 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFQTSCEITGFGK 319

QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 320 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 379

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 380 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 430

```

```

RESULT 9
US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A

```

```

; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48

```

```

Query Match          99.08; Score 2277; DB 2; Length 411;
Best Local Similarity 99.38; Pred. No. 2.6e-190;
Matches 408; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKKEGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWSNATVILQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWSNATVILQOTYHAHRSDALQGLGKHNYCRNPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRPKIIGGFTTIENQFWFAAIYRRH 180
DB 121 PLVQECMVHDCADGKPPSPPEELKFCQCGKTLRPRPKIIGGFTTIENQFWFAAIYRRH 180

QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 240
DB 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENLI 240

QY 241 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFQTSCEITGFGK 300
DB 241 LHKDYSADTLAHNDIALKIRKEGRCAPSPRTIQTICLPSMYNDPQFQTSCEITGFGK 300

QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411

```

```

RESULT 10
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEEDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```


APPLICANT: SAUNDERS, DEREK J.
TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
TITLE OF INVENTION: INHIBITING EFFECT
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N. W. Suite 700
CITY: Washington, D.C.
COUNTRY: U.S.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/720,012
FILING DATE: 27-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,741
FILING DATE: 20-JUL-1993
APPLICATION NUMBER: DE P43 23 754.1
FILING DATE: 15-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/41345
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)628-8800
TELEFAX: (202)628-8844
INFORMATION FOR SEQ ID NO: 83:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-720-012-83

Query Match 87.9%; Score 2022; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 3.5e-168;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHSDALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHSDALQLGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEELKFCQGQKTLRPRFKIIGGEFTT 166
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEELKFCQGQKTLRPRFKIIGGEFTT 120
QY 167 TENQPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 226
Db 121 TENQPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 180
QY 227 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 286
Db 181 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 240
QY 287 PQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHRECQQPHYYSVETTTKMLCAADPQ 346
Db 241 PQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHRECQQPHYYSVETTTKMLCAADPQ 300
QY 347 WKTDSQCGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGYTRVSHFLPWIRSHTKEE 406
Db 301 WKTDSQCGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGYTRVSHFLPWIRSHTKEE 360
QY 407 NGLAL 411
Db 361 NGLAL 365

RESULT 13

US-08-560-098A-44
Sequence 44, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WNEENDT, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins Having Fibrinolytic and
TITLE OF INVENTION: Coagulation-Inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-44

Query Match 87.9%; Score 2022; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 3.8e-168;
Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHSDALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOTYHAHSDALQLGLGKHNYCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEELKFCQGQKTLRPRFKIIGGEFTT 166
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSPPEELKFCQGQKTLRPRFKIIGGEFTT 121
QY 167 TENQPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 226
Db 122 TENQPFPAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNS 181
QY 227 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 286
Db 182 NTQGMKEFEVENLILHKDYSADTLAHNDIALLKIRSKEGRCQAQPSRTIQTICLPSMYND 241
QY 287 PQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHRECQQPHYYSVETTTKMLCAADPQ 346
Db 242 PQFGTSCBITGFGKENSTDYLYPEQLKMTVVKLISHRECQQPHYYSVETTTKMLCAADPQ 301
QY 347 WKTDSQCGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGYTRVSHFLPWIRSHTKEE 406
Db 302 WKTDSQCGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGYTRVSHFLPWIRSHTKEE 361
QY 407 NGLAL 411

Db 362 NGLAL 366

|||||

RESULT 14

US-08-967-024C-24

; Sequence 24, Application US/08967024C

; Patent No. 6133011

; GENERAL INFORMATION:

; APPLICANT: WNEEDT, Stephan

; APPLICANT: STEFFENS, Gerd Josef

; APPLICANT: JANOSHA, Elke

; APPLICANT: HEINZEL-WIELAND, Regina

; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/967,024C

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 42 665.8

; FILING DATE: 30-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 148/42444

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 393 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-967-024C-24

Query Match 87.9%; Score 2022; DB 3; Length 393;

Best Local Similarity 100.0%; Pred. No. 3.8e-168;

Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQYHAHRSDALQLGLGKHNCRNP 106

Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQYHAHRSDALQLGLGKHNCRNP 61

Qy 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKIIGSEFTT 166

Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKIIGSEFTT 121

Qy 167 IENQFWFAAIIYRRHGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 226

Db 122 IENQFWFAAIIYRRHGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 181

Qy 227 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 241

Db 287 PQFGTSCIEITGFGKENSVDYLYPEOLKMTVVKLISHRECQPHYVGSVTTKMLCAADPQ 346

Db 242 PQFGTSCIEITGFGKENSVDYLYPEOLKMTVVKLISHRECQPHYVGSVTTKMLCAADPQ 301

Qy 347 WKTDSCQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHYKEE 406

Db 302 WKTDSCQDSGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHYKEE 361

Qy 407 NGLAL 411

Db 362 NGLAL 366

RESULT 15

US-08-967-024C-25

; Sequence 25, Application US/08967024C

; Patent No. 6133011

; GENERAL INFORMATION:

; APPLICANT: WNEEDT, Stephan

; APPLICANT: STEFFENS, Gerd Josef

; APPLICANT: JANOSHA, Elke

; APPLICANT: HEINZEL-WIELAND, Regina

; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan

; STREET: 1200 G Street, N.W., Suite 700

; CITY: Washington

; STATE: DC

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/967,024C

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: P 44 42 665.8

; FILING DATE: 30-NOV-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: EVANS, Joseph D.

; REGISTRATION NUMBER: 26,269

; REFERENCE/DOCKET NUMBER: 148/42444

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 628-8800

; TELEFAX: (202) 628-8844

; INFORMATION FOR SEQ ID NO: 25:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 393 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-967-024C-25

Query Match 87.9%; Score 2022; DB 3; Length 393;

Best Local Similarity 100.0%; Pred. No. 3.8e-168;

Matches 365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQYHAHRSDALQLGLGKHNCRNP 106

Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQOQYHAHRSDALQLGLGKHNCRNP 61

Qy 107 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKIIGSEFTT 166

Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGKKPSSPEELKFCQCKTLRPRFKIIGSEFTT 121

Qy 167 IENQFWFAAIIYRRHGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 226

Db 122 IENQFWFAAIIYRRHGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNS 181

Qy 227 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 286

Db 182 NTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYND 241

QY 287 POFGTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQ 346
Db 242 POFGTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLCAADPQ 301
QY 347 WKTDSQQGSDGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHKDE 406
Db 302 WKTDSQQGSDGGLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHKDE 361
QY 407 NGLAL 411
Db 362 NGLAL 366

RESULT 16
US-08-560-098A-45
; Sequence 45, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEINDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gard Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-45

Query Match 65.5%; Score 1507.5; DB 2; Length 306;
Best Local Similarity 90.6%; Pred. No. 1.7e-123;
Matches 281; Conservative 3; Mismatches 9; Indels 17; Gaps 2;
QY 103 RNPONRRP-WCYQVGLKPLVQECWVHDCADGKKPSPPEELKFCQCKTLRPRFKIIG 161
Db 13 RNPNDKYEPFW-----EDEKGMSPPEELKFCQCKTLRPRFKIIG 56
QY 162 GEFTTIENQPMFAAIYRRHGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 221
Db 57 GEFTTIENQPMFAAIYRRHGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 116
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSGRCAQPSRTIQTICLP 281

Db 117 SRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSGRCAQPSRTIQTICLP 176
QY 282 SMYNDQFQTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLC 341
Db 177 SMYNDQFQTSCEITGFGKENSVDLYPEQLKMTVVKLISHRECQPHYYGSEVTTKMLC 236
QY 342 AADPQKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRS 401
Db 237 AADPQKTDSCQDGGPLVCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRS 296
QY 402 HTKEENGLAL 411
Db 297 HTKEENGLAL 306

RESULT 17
US-08-560-098A-46
; Sequence 46, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEINDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gard Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-46

Query Match 65.5%; Score 1507.5; DB 2; Length 331;
Best Local Similarity 90.6%; Pred. No. 1.9e-123;
Matches 281; Conservative 3; Mismatches 9; Indels 17; Gaps 2;
QY 103 RNPONRRP-WCYQVGLKPLVQECWVHDCADGKKPSPPEELKFCQCKTLRPRFKIIG 161
Db 13 RNPNDKYEPFW-----EDEKGMSPPEELKFCQCKTLRPRFKIIG 56
QY 162 GEFTTIENQPMFAAIYRRHGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 221
Db 57 GEFTTIENQPMFAAIYRRHGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVLGR 116
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALLKIRSGRCAQPSRTIQTICLP 281

Db 117 SRLNSNTQGMKFVENILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQTICLP 176
QY 282 SMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGVSEVTTKMLC 341
Db 177 SMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGVSEVTTKMLC 236
QY 342 AADPQWKTDCQSGSGPLVCSLOGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRS 401
Db 237 AADPQWKTDCQSGSGPLVCSLOGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPWIRS 296
QY 402 HTKEENGLAL 411
Db 297 HTKEENGLAL 306

RESULT 18
US-08-944-483-73
; Sequence 73, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-73

Query Match 60.1%; Score 1382; DB 3; Length 253;
Best Local Similarity 100.0%; Pred. No. 1.2e-112;
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 IIGGEFTTIENQWFAAIYRRHRGGSVTVVCGSLISPCWVISAHCFIDYPKKEDIYV 218
Db 1 IIGGEFTTIENQWFAAIYRRHRGGSVTVVCGSLISPCWVISAHCFIDYPKKEDIYV 60
QY 219 LGRSLNSNTQGMKFVENILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQT 278
Db 61 LGRSLNSNTQGMKFVENILHKDYSADTLAHNDIALLKIRSGRCAQPSRTIQT 120
QY 279 CLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGVSEVTTK 338
Db 121 CLPSMYNDPQGTSCITGFGKENSTDYLYPEQLKMTVVKLIHRECCQPHYGVSEVTTK 180
QY 339 MLCAADPQWKTDCQSGSGPLVCSLOGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPW 398
Db 181 MLCAADPQWKTDCQSGSGPLVCSLOGRMTLTGIVSWGRGKALKDKPGVYTRVSHFLPW 240
QY 399 IRSHTKEENGLAL 411
Db 241 IRSHTKEENGLAL 253

RESULT 19
US-08-560-098A-49
; Sequence 49, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEENDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-49

Query Match 59.7%; Score 1374; DB 2; Length 254;
Best Local Similarity 99.2%; Pred. No. 5.8e-112;
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 158 KIIGGEFTTIENQWFAAIYRRHRGGSVTVVCGSLISPCWVISAHCFIDYPKKEDIYV 217
Db 1 KIIGGEFTTIENQWFAAIYRRHRGGSVTVVCGSLISPCWVISAHCFIDYPKKEDIYV 60

Y YLGRSLNNTQGMKFEVENILHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQT 277
Y YLGRSLNNTQGMKFEVENILHKDYSADTLAHHNDIALKIRSKGRCQAPSTIQT 120
Y ICLPSMYNDPQFQTSCEITGFGKENSTLYPEQLKMTVVVKLIHRECOQPHYGEVTT 337
Y ICLPSMYNDPQFQTSCEITGFGKENSTLYPEQLKMTVVVKLIHRECOQPHYGEVTT 180
Y KMLCAADPQWKTSCOGDSGGLVCSLOGRMTLTGLVSGRCALKDKGVTVTRVSHFLP 397
Y KMLCAADPQWKTSCOGDSGGLVCSLOGRMTLTGLVSGRCALKDKGVTVTRVSHFLP 240
Y WIRSHKTEENGLAL 411
Y WIRSHKTEENGLVL 254
RESULT 20
US-09-101-272G-73
Sequence 73, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101.272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: Patent in version 3.1
SEQ ID NO 73
LENGTH: 200
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73
Query Match 45.2%; Score 1041; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.2e-83;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRFKIIGGFTTIENQPWFAAIYRRH 180
DB 141 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRFKIIGGFTTIENQPWFAAIYRRH 200
RESULT 21
US-08-142-590B-25
Sequence 25, Application US/08142590B
Patent No. 6120765
GENERAL INFORMATION:
APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and GOETINCK,
TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/142,590B
FILING DATE: 25-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,318
FILING DATE: 02-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Paul L.
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: MGP-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-142-590B-25
Query Match 38.4%; Score 883; DB 3; Length 157;
Best Local Similarity 96.8%; Pred. No. 1.8e-69;
Matches 152; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGTCTVSNKYFSNIHWCNCPKFKGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRF 157
DB 121 PLVQECMVHDCADGKSPPEELKFCQGGOKTLRPRF 157
RESULT 22
US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997


```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/01025A
; FILING DATE: 19910214
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/486,657
; FILING DATE: 1 March 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hasak, Janet E.
; REGISTRATION NUMBER: 28,616
; REFERENCE/DOCKET NUMBER: 454P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/266-1896
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; PCT-US91-01025A-2

Query Match
Best Local Similarity 37.6%; Score 864.5; DB 5; Length 527;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPGEGAGKCCEDTATCYEQGISY 100
QY 59 RKGASTDTWGRCLPWNATVILQOYTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWTAESGAECTWNWSSALAKPYSGRRPDAIRLGLGNHNYCRNPNDRDSPKCVFKA 160
QY 119 LKPLVQECMVHDCADG-----KGPS 138
Db 161 GYSEFCSTPACSEGNSDCYFNGSAVGRGTHSLTESGASCLPWNMSMILIGVYTAQNP 220
QY 139 -----SPEELK-----PQCG-OKTLRPRFKIIGGE 163
Db 221 AQALGLGKHNYCRNPDGDAKPCWCHLVKNRRLTWECVDPSCSTGLRQVYQFRIKGL 280
QY 164 FTTIENQWFAIYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSR 282
Db 281 FADIASHPQAAIFAKHRRSPGERFLCGGILISSCWILSAACFCQERPPPHLTVILGRT 340
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLLKIRSKEGRCQAQPSRTIOTICLPS 398
Db 341 YRVVPEEEQKFEVEKYIVHKEFDDT--YNDIALQLKSDSSRCAQESSVVRVCLPP 398
QY 283 MYNDPQFTSCBITGFGKENSTDYLPOLKMTVVKLISHRECCQPHYYGSEVTTMLCA 342
Db 399 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 458
QY 343 AD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRCALDKPGVYTVSHFL 396
Db 459 GDRSGGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKDPVGYTVKVTNYL 518
QY 397 PWIRSHTK 404
Db 519 DWIRDNMR 526

RESULT 25
5185259-8
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.

```

```

; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1992
; SEQ ID NO: 8:
; LENGTH: 527
5185259-8

Query Match
Best Local Similarity 37.6%; Score 864.5; DB 6; Length 527;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQALYFSDF-VQCPGEGAGKCCEDTATCYEQGISY 100
QY 59 RKGASTDTWGRCLPWNATVILQOYTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 101 RGTWTAESGAECTWNWSSALAKPYSGRRPDAIRLGLGNHNYCRNPNDRDSPKCVFKA 160
QY 119 LKPLVQECMVHDCADG-----KGPS 138
Db 161 GYSEFCSTPACSEGNSDCYFNGSAVGRGTHSLTESGASCLPWNMSMILIGVYTAQNP 220
QY 139 -----SPEELK-----PQCG-OKTLRPRFKIIGGE 163
Db 221 AQALGLGKHNYCRNPDGDAKPCWCHLVKNRRLTWECVDPSCSTGLRQVYQFRIKGL 280
QY 164 FTTIENQWFAIYRRH-RGGSVTVVCGSLISPCWVISATHCFIDYPKKEDYIVYLGSR 282
Db 281 FADIASHPQAAIFAKHRRSPGERFLCGGILISSCWILSAACFCQERPPPHLTVILGRT 340
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLLKIRSKEGRCQAQPSRTIOTICLPS 398
Db 341 YRVVPEEEQKFEVEKYIVHKEFDDT--YNDIALQLKSDSSRCAQESSVVRVCLPP 398
QY 283 MYNDPQFTSCBITGFGKENSTDYLPOLKMTVVKLISHRECCQPHYYGSEVTTMLCA 342
Db 399 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 458
QY 343 AD-----POWKT-DSCQDGGPLVCSLQGRMTLTGIVSWGRCALDKPGVYTVSHFL 396
Db 459 GDRSGGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKDPVGYTVKVTNYL 518
QY 397 PWIRSHTK 404
Db 519 DWIRDNMR 526

RESULT 26
US-08-811-949-43
; Sequence 43, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67

```

RESULT 27

US-08-560-098A-50
; Sequence 50, Application US/08560098A
; Patent No. 5978841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D

```

, REFERENCE/DOCKET NUMBER: 148/42448
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: (202) 628-8800
, TELEFAX: (202) 628-8844
, INFORMATION FOR SEQ ID NO: 50:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 562 amino acids
, TYPE: amino acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: protein
, US-08-560C-098A-50

```

Query Match	37.6%	Score	864.5	DB 2	Length	562
Best Local Similarity	38.1%	Pred. No.	3.6e-67			
Matches	186	Conservative	56	Mismatches	157	Indels
						Gaps
						11

Qy	3	ELHOVP-SNCD---CLNGGTCVSNKYESNIHWCNCPKFKFGQHOCEIDKSKTCYEGNHGFY	58
Db	77	QCHSVFVKVSCSEPRCFNGGTCQALYSDP-VQCQPEGFAGKCEIDTRATCYDQGGISY	135
Qy	59	RGKASTDTMRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRRRPWCYQVG	118
Db	136	RGTWSTAESGAECTWNSSALAAQKPYSGRRPDAIRLGLGNHNYCRNPDORSKPCYVFKA	195
Qy	119	LKPLVQSCMVHDCADG-----KPKS	138
Db	196	GKYSSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGKVYTAQNP	255
Qy	139	-----SPPEELK-----FQCG-QKTLPREFKIIGGE	163
Db	256	AQALGLGKHNYCRNPDGDAKFWCHLVKNRRLTWECYDVPSCSTCGLRQYSQPQFRIKGG	315
Qy	164	FTTIENQPFALYRRH-RGGSVTVYCGGSLISPCWVISATHCFIDYPKXEDIYVLGRS	222
Db	316	FADIASHPQALFAKHRRSPGERFLCCGLHISSCWILSAACHQERFPPHLLTVILGR	375
Qy	223	RLNSNTGCKMFVENLILHKDYSADTLAHHNDIALIKRKEGRCAQPSRTIQTICLPS	282
Db	376	YRVVPGEEEOKFVEKYIVHKEFDDT--YDNDIALLOLKSDSRCAOESSVWATCLPP	433

[illegible]

Db 376 YRVPGEEEOKFVEKYIVHKEFFDDT--YDNDIALLOLKSDSSRCAOESSVVRTVCLPF 433

QY 283 MYNDPQGTSCBTGRKENSTDYLYPEQLKMTVVKLI SHRECOQPHYYGSEVTTKMLCA 342
Db 434 ADLQLPDWTCELSGKGKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493
QY 343 AD-----PQWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGALKDKFQVYTRYSHFL 396
Db 494 GDTRSGGPQANLHDAQCQDGGGLVCLNDGRMTLVGLIISWGLGCGQKDVGVYTKVTNYL 553
QY 397 PWIRSHTK 404
Db 554 DWIRDNR 561
RESULT 28
US-08-883-795A-38
; Sequence 36, Application US/08893795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcuve, Genevieve
; APPLICANT: Awang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERSKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein (tpa)
US-08-883-795A-38
Query Match 37.6%; Score 864.5; DB 2; Length 562;
Best Local Similarity 38.1%; Pred. No. 3.6e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSBRFCFNGGTCQQALYFSDF-VQCQPEGFAGKCEIDTRATCYEDQG1SY 135
QY 59 RGKASDTWGRCLPWN SATVLQOYTHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDAIRLGLGNHNYCRPNDRSDKPCWCVFKA 195
QY 119 LKPLVQECMVHDCADG-----KGPS 138
Db 196 GKYSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWN SMILLIKVYTAQNS 255
QY 139 -----SPEELK-----PQCG-QKTLRPRFKIIGE 163
Db 256 AQALGLGKHNYCRNP DGDAPKPVCHLVKNRRLTW EYCDVPSCSTCGLRQYSQQFRIKGG 315
QY 164 FTTIENQPFAPAIYRRH-RGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 222
Db 316 FADIASHPMQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 375
QY 223 RLNSNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSKGRCACQSPRTIOTICLPS 282

Db 256 AQALGLGKHNYCRNP DGDAPKPVCHLVKNRRLTW EYCDVPSCSTCGLRQYSQQFRIKGG 315
QY 164 FTTIENQPFAPAIYRRH-RGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 222
Db 316 FADIASHPMQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 375
QY 223 RLNSNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSKGRCACQSPRTIOTICLPS 282
Db 376 YRVVPEEEQKFEVEKYIVHKEFDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433
QY 283 MYNDPQGTSCBTGRKENSTDYLYPEQLKMTVVKLI SHRECOQPHYYGSEVTTKMLCA 342
Db 434 ADLQLPDWTCELSGKGKHEALSPFYSERLKEAHRVLYPSSRCTSQHLLNRTVTDNMLCA 493
QY 343 AD-----PQWKT-DSQCGSGGGLVCSLQGRMTLTGIVSWGRGALKDKFQVYTRYSHFL 396
Db 494 GDTRSGGPQANLHDAQCQDGGGLVCLNDGRMTLVGLIISWGLGCGQKDVGVYTKVTNYL 553
QY 397 PWIRSHTK 404
Db 554 DWIRDNR 561
RESULT 29
5185259-3
; Patent No. 5185259
; APPLICANT: GOEDDEL, DAVID V.; KOHR, WILLIAM J.; PENNICA, DIANE;
; VEHAR, GORDON A.
; TITLE OF INVENTION: TRUNCATED HUMAN TISSUE PLASMINOGEN
; ACTIVATOR
; NUMBER OF SEQUENCES: 15
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/489,855
; FILING DATE: 02-MAR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 12,694
; FILING DATE: 09-FEB-1987
; APPLICATION NUMBER: 483,052
; FILING DATE: 07-APR-1983
; APPLICATION NUMBER: 398,003
; FILING DATE: 14-JUL-1982
; APPLICATION NUMBER: 374,860
; FILING DATE: 05-MAY-1982
; SEQ ID NO: 3;
; LENGTH: 562
5185259-3
Query Match 37.6%; Score 864.5; DB 6; Length 562;
Best Local Similarity 38.1%; Pred. No. 3.6e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSBRFCFNGGTCQQALYFSDF-VQCQPEGFAGKCEIDTRATCYEDQG1SY 135
QY 59 RGKASDTWGRCLPWN SATVLQOYTHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAKPYSGRRPDAIRLGLGNHNYCRPNDRSDKPCWCVFKA 195
QY 119 LKPLVQECMVHDCADG-----KGPS 138
Db 196 GKYSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWN SMILLIKVYTAQNS 255
QY 139 -----SPEELK-----PQCG-QKTLRPRFKIIGE 163
Db 256 AQALGLGKHNYCRNP DGDAPKPVCHLVKNRRLTW EYCDVPSCSTCGLRQYSQQFRIKGG 315
QY 164 FTTIENQPFAPAIYRRH-RGGSVTVYCGSLISPCWVISATHCFIDYPKKEDYIVVLGRS 222
Db 316 FADIASHPMQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPFPHLTVILGRT 375
QY 223 RLNSNTQGMKFEVENLIHKDYSADTLAHNDIALLKIRSKGRCACQSPRTIOTICLPS 282

376	YRVVPEBBQBKFEVEKVIIVHKEFPDDDT--YONDYALLQIKSDSSCAQESSVVVRIVCI	PP	433
283	MYNDPQGTSCETITGFGKENSTDYLPQLKMTVVVKLI	SHRECCQPHFYGSEVTTTMLCA	342
434	ADLQLPDWTECELSYGKHEALSPFYSERLKEAHVRLPSPSRCTSQHLLNLRIVTDNMLCA		493
343	AD-----PQWKT-DSQCQDSSGGPLVCSLQGRMTLTGIVSWGRCALKKQPGVYTRVSHPL		396
494	GDTRSGGPGQANLHDACQDSSGGPLVCLNDGRMTLVGLIISWLGCGCKQKQPGVYTKVTNYL		553
397	PWIRSHTK	404	
554	DWIRDNMR	561	

RESULT 30

RESULI 30
5200340-2

5200340-2
: Patent No 5200340

APPLICANT: EOSTER DONALD C .MUTVYTHILL. ETTLEEN P .O'HARA

APPLICANT: FOSTER, DONALD C.; MULVHILL,
PATRICK I.; PINCEI, KURT; YOSHITAKE SHINJI

;PATRICK J.;PINGEL, KURT;YOSHITAKE, SHINJI
: TITIE OE INTENTION: THROMBIN-ACTIVATED TISSUE PLASMINOGEN

ACTIVATORS ; TITLE OF INVENTION: THREE

ACTIVATORS	NUMBER OF SEQUENCES
1	34

NUMBER OF SEQUENCES: 34

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 100-10753410

```

APPLICATION NUMBER: US/

FILING DATE: 22-MAY-1987

```

;SEQ ID NO:2:

```

LET
;

Query Match	37.6%	Score	854.5	DB	6	Length	562
Best Local Similarity	38.1%	Pred. No.	3.6e-67				
Matches	186	Conservative	56	Mismatches	157	Indels	89
						Gaps	11

QY 3 ELHQVP-SNCD--CLNGGTCSNKFYSNIHWCNCPKFGGQHC EIDKSKTCYEGNGHFY 58

db 77 OCHSVPVKSCSEPRCFNGGTCCOALYFSDF-VCCPGEFAGKCCCEIDTRATCYEDGISY 135

59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVG 118

[illegible]

Q7
110 Y YBY WAFBOMWUPCZND

WYD C 138

QY	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128
	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128

196 GKYSS EFCST PACSE GNSDCYFGNGS AYRGTHSLT ES GASCLPWNS MILIGKVYTAQNPS 255

QY 139 -----SPPEELK-----FQCG-QKTLRPFKIIGGE 163

Db 256 AQAALGLGKHNYCRNPDGDAKPWCHVLKNRRLLTWEYCDVPSCSTCGLRQYSQPOFRIKGG 315

164 EMTTBNQDWEAATYDPH-DCCGUTVUCCGCTGGCMTGAMICEEDVDEKPEEDYINVTGDC

QY		:											:					:																																
104	F	Y	I	E	N	Q	F	F	A	T	K	R	H	-	R	G	S	V	I	V	C	G	S	L	T	S	F	C	W	V	I	S	A	H	C	F	D	I	F	N	E	D	I	V	I	D	G	R	S	222

Db 316 FADIAHPWQAAIFAKHRRSPGERFLCGGILISSCWILSAAHCFQERFPPHHLTVILGRT 375

QY 223 RLNSNTQGMKFVENLILHKDYSADTLAHNDIALLKIRSKEGRCAPSRTIQTICLPS 282

```

Db
376 YRVPGEEQKFEVEKYIVHKEFDDT--YDNDIALLOLKSDSSRCAOESSVVRTVCLPP 433

```

[illegible]

QY 283 MINDFQGISCEITGFGRENSIDLILPEQLRMIVVATLISHRECQPHHYGSEVTRMLCA 342

Db 434 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHVRLYPSSRCTSQHLLNRTVTDNMLCA 493

QY 343 AD-----PQWKT-DSCQDSDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396

Search completed: December 3, 2003, 14:45:34
Job time : 23.2586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 64.7422 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNELHQVPSNCCLNGTCTV.....VSHFLPWIRSHKTEGLAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2301	100.0	411	6 AAP50871	Sequence encoded b
2	2301	100.0	411	11 AAR06244	Urokinase precursor
3	2301	100.0	411	16 AAR62991	Pro-urokinase. Ho
4	2301	100.0	411	21 AAY92836	Urokinase plasmino
5	2301	100.0	411	23 AAE16544	Human urokinase-ty
6	2301	100.0	431	6 AAP50114	Sequence encoded b
7	2301	100.0	431	7 AAE60783	Human urokinase.
8	2301	100.0	431	8 AAP70258	Sequence of human
9	2301	100.0	431	9 AAP80430	Deduced AA sequenc

10	2301	100.0	431	9 AAP81204	Pro-urokinase with
11	2301	100.0	431	10 AAP92119	Natural human prou
12	2301	100.0	431	11 AAR04253	Human pro-urokinas
13	2301	100.0	431	11 AAR07112	Human pro-Urokinas
14	2301	100.0	431	15 AAR47903	Pro-urokinase deri
15	2301	100.0	431	15 AAR63141	Full length human
16	2301	100.0	431	21 AAY99591	Human plasminogen
17	2301	100.0	431	21 AAY50869	Human urokinase pr
18	2301	100.0	431	22 AAB84605	Amino acid sequenc
19	2301	100.0	431	23 AAG79460	sc-UPA. Homo sapi
20	2301	100.0	431	23 AAU99228	Human plasminogen
21	2301	100.0	431	23 ABE17128	Human uPA protein.
22	2301	100.0	431	24 ABU56547	Lung cancer-associ
23	2301	100.0	431	24 ABU56708	Lung cancer-associ
24	2301	100.0	431	24 ABU11076	Human urokinase pl
25	2301	100.0	434	13 AAR20537	Amidated deriv. of
26	2301	100.0	434	13 AAR20538	Amidated deriv. of
27	2301	100.0	436	13 AAR20536	Amidated deriv. of
28	2298	99.9	430	18 AAW24578	Inhibitor resistan
29	2298	99.9	431	10 AAP91886	Sequence of prouro
30	2298	99.9	431	10 AAP94764	Non-glycosylated p
31	2297	99.8	411	14 AAR34584	Mutant human prou
32	2297	99.8	431	23 AAU99230	Human plasminogen
33	2296	99.8	411	10 AAP96146	Sequence encoded b
34	2296	99.8	411	17 AAR92526	Pro-urokinase. Ho
35	2296	99.8	411	22 AAB74797	Prourokinase prote
36	2296	99.8	431	23 AAU99229	Human plasminogen
37	2295	99.7	411	16 AAR62992	Pro-urokinase muta
38	2295	99.7	411	16 AAR62993	Pro-urokinase muta
39	2295	99.7	411	16 AAR62998	Pro-urokinase muta
40	2295	99.7	411	16 AAR62999	Pro-urokinase muta
41	2293	99.7	411	16 AAR63000	Pro-urokinase muta
42	2293	99.7	411	20 AAY39343	Human pro-urokinas
43	2293	99.7	411	20 AAY42284	Human pro-urokinas
44	2293	99.7	411	22 AAB20489	Human pro-urokinas
45	2293	99.7	412	12 AAR10334	Recombinant single

ALIGNMENTS

RESULT 1

AAP50871 ID AAP50871 standard; protein; 411 AA.

XX AC AAP50871;

XX DT 30-NOV-1991 (first entry)

XX DE Sequence encoded by cDNA sequence for human urokinase zymogen

DE (Japanese Patent Application No.37119/84).

XX KW Thrombolytic agent; plasminogen activator activity; fibrin affinity;

KW enzyme.

XX OS Homo sapiens.

XX FH Key

FT Cleavage-site

FT Location/Qualifiers

FT 158..159

FT /note= "potential cleavage site which generates

FT the two-chain form from the zymogen"

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX EP139447-A.

```
PD 02-MAY-1985.
XX
PF 07-SEP-1984; 84EP-0306117.
XX
PR 17-OCT-1983; 83JP-0195051.
PR 13-SEP-1983; 83JP-0170354.
XX
PA (GREC ) GREEN CROSS CORP.
XX
PI Kasai S, Arimura H, Mori K, Nishida M, Suyama T;
XX WPI; 1985-106530/18.
XX
PT New urokinase zymogen - useful as thrombolytic agent
XX
PS Disclosure; Page 12; 30pp; English.
XX
CC Zymogen AAP50871 is the inactive precursor form of human urokinase.
CC Urokinase zymogen is cleaved into the two-chain form composed of
CC characteristic urokinase H (molecular wt. of 30,000) and L (molecular
CC wt. of 20,000) chains when treated with catalytic amounts of plasmin.
CC The patentors claim a new urokinase zymogen which has mol. wt. ca.
CC 50,000, a single chain molecular structure, and selective affinity
CC for fibrin. It is a thrombolytic agent which manifests its
CC plasminogen activator activity on cleavage by proteolytic enzymes
CC (e.g. plasmin) and has higher affinity for fibrin than known forms
CC of urokinase.
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 2301; DB 6; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.4e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPWCYVOVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPWCYVOVGLK 120
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
DB 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENLI 240
DB 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENLI 240
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQITICLPMSYNDPQGTSCETIGFGK 300
DB 241 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQITICLPMSYNDPQGTSCETIGFGK 300
QY 301 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGL 360
DB 301 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGL 360
QY 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTVRSHFLPWIRSHTKENGIAL 411
DB 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTVRSHFLPWIRSHTKENGIAL 411
RESULT 2
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX
AC AAR06244;
XX
XX 07-DEC-1990 (first entry)
DT
DE Urokinase precursor protein.
XX
```

```
KU Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
KW myocardial infarction;
OS Homo sapiens.
XX
PN EP380334-A.
XX
PD 01-AUG-1990.
XX
PF 25-JAN-1990; 90EP-0300772.
XX
PR 17-MAY-1989; 89JP-0121405.
PR 27-JAN-1989; 89JP-0016406.
XX
PA (GREC ) GREEN CROSS CORP.
XX
PI Matsuda H, Ueda Y, Tamanouchi K;
XX WPI; 1990-233117/31.
XX
PT Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bioavailability and improved activity
XX
PS Claim 3; Fig 1; lipp; English.
XX
CC By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 2301; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.4e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPWCYVOVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPWCYVOVGLK 120
QY 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
DB 121 PLVQECMWHDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPWFAAIYRRH 180
QY 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENLI 240
DB 181 RGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVEVENLI 240
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQITICLPMSYNDPQGTSCETIGFGK 300
DB 241 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQITICLPMSYNDPQGTSCETIGFGK 300
QY 301 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGL 360
DB 301 ENSTDYLYPEQLKMTVVVKLI SHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDGGGGL 360
QY 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTVRSHFLPWIRSHTKENGIAL 411
DB 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTVRSHFLPWIRSHTKENGIAL 411
RESULT 3
AAR62991
ID AAR62991 standard; protein; 411 AA.
XX
AC AAR62991;
XX
```


[illegible]

Db	181	RGGSVTVYCGGSLISPCWVIGATHCFIDYIPKKEGVIVYIGRRLNSNTGEMKFEVENLI	241	
Qy	241	LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTTIQTICLPSMYNDPQFGTSCEITGFGK	300	
Db	241	LHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTTIQTICLPSMYNDPQFGTSCEITGFGK	300	
Qy	301	ENSTDYLYPEQLKMTVVKLIISHRECCQPHYGVSEVTTTMLCAADPQWKTDSCQDSGGPL	360	
Db	301	ENSTDYLYPEQLKMTVVKLIISHRECCQPHYGVSEVTTTMLCAADPQWKTDSCQDSGGPL	360	
Qy	361	VCSLQGRMTLNGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKBENGIAL	411	
Db	361	VCSLQGRMTLNGIVSWGRCALKDKPGVYTVRSHFLPWIRSHTKBENGIAL	411	
RESULT 4				
AA92836				
ID	AA92836 standard; Protein; 411 AA.			
XX	AC	AA92836;		
XX	XX			
DT	29-AUG-2000 (first entry)			
XX				
DE	Urokinase plasminogen activator (uPA).			
XX				
KW	N-terminal; pro-uPA; urokinase plasminogen activator; receptor; uPAR;			
KW	anti-cancer; anti-metastatic; anti-proliferative; anti-atherosclerotic;			
KW	anti-thrombotic; anti-angiogenic; anti-inflammatory; anti-arthritis;			
KW	anti-fibrotic; apoptotic; vasotropic; anti-diabetic; ophthalmological;			
KW	thrombolytic.			
XX				
OS	Homo sapiens.			
XX				
PH	Key	Location/Qualifiers		
FT	Disulfide-bond	11..19		
FT	Disulfide-bond	13..31		
FT	Disulfide-bond	33..42		
FT	Disulfide-bond	50..131		
FT	Disulfide-bond	71..113		
FT	Disulfide-bond	102..126		
FT	Disulfide-bond	148..279		
FT	Disulfide-bond	189..205		
FT	Disulfide-bond	197..268		
FT	Disulfide-bond	293..362		
FT	Disulfide-bond	325..341		
FT	Disulfide-bond	352..380		
XX				
PN	WO200026353-A1.			
XX				
PD	11-MAY-2000.			
XX				
PF	28-OCT-1999; 99WO-US25210.			
XX				
PR	29-OCT-1998; 99US-0181816.			
XX				
PA	(ANGS-) ANGSTROM PHARM INC.			
XX				
PI	Mazar AP, Jones TR;			
XX				
DR	WPI; 2000-365605/31.			
XX				
PT	New cyclic peptide, useful for treatment or diagnosis of e.g. tumors			
PT	and other diseases involving cell proliferation or migration, targets			
PT	the urokinase plasminogen activator receptor			
XX				
PS	Disclosure; Fig 1; 93pp; English.			
XX				
CC	The present sequence shows the wild-type urokinase plasminogen activator			
CC	(uPA). Cyclic peptides based on the amino acids residues 20-30 (the			
CC	receptor-binding region) of uPA are claimed. These cyclic peptides target			
CC	the uPA receptor (uPAR), allowing therapeutic or diagnostic agents to be			
CC	delivered to uPAR-expressing cells. The cyclic peptides are used,			
CC	optionally when linked to a therapeutic agent, to inhibit migration.			

XX AC AAP50114;
XX AC
XX 27-SEP-1991 (first entry)
XX
XX Sequence encoded by the signal sequence and noncoding region of the
XX DE pro-UK structural gene (Sequence II).
XX
XX Enzyme; thrombosis therapy; embolic disease;
XX KW single-chain pro-urokinase.
XX KW
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
XX FT Domain 21..177
XX FT /label= A chain
XX FT Domain 179..431
XX FT /label= B chain
XX FT Peptide 1..20
XX FT /label= signal peptide
XX
XX EP154272-A.
XX PN
XX 11-SEP-1985.
XX PD
XX 23-FEB-1985; 85EP-0102031.
XX PF
XX 31-JAN-1985; 85JP-0017969.
XX PR
XX 27-FEB-1984; 84JP-0037119.
XX PR
XX (GREC) GREEN CROSS CORP.
XX PA
XX Hiramatsu R, Kaneda T, Nagai M, Arimura H, Nishida M;
XX PI Suyama T;
XX PI
XX WPI: 1985-224693/37.
XX DR N-PSDB; AAN50138.
XX
XX Glycosylated single-chain pro-urokinase - prep'd. by cultivating
XX FT animal cells transformed by DNA prep'd. from m RNA
XX FT
XX Disclosure; Page 8-10; 64pp; English.
XX PS
XX The inventors claim a method of producing single-chain
XX CC pro-urokinase by using as template, mRNA obtd. from cells of an
XX CC established human kidney-derived cell line. The urokinase is used to
XX CC treat thrombosis and embolic diseases as well as in the treatment of
XX CC diseases in combination with anticancer agents.
XX CC
XX SQ Sequence 431 AA;
Query Match 100.0%; Score 2301; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 5,7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYVQGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYVQGLK 140
QY 121 PLVQECMVHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
DB 141 PLVQECMVHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 200
QY 181 RGSVTVYCGSLISPCWISATHCFFIDYPKEDYIVYIGRSLNSNTQGMKFEVNI 240
DB 201 RGSVTVYCGSLISPCWISATHCFFIDYPKEDYIVYIGRSLNSNTQGMKFEVNI 260
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQITCLPSMYNDPQFSGTSCITGFGK 300
|||||

DB 261 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQITCLPSMYNDPQFSGTSCITGFGK 320
QY 301 ENSTDYLYPEQLKMTYVVKLIISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSGGPL 360
DB 321 ENSTDYLYPEQLKMTYVVKLIISHRECOQPHYGSEVTTKMLCAADPOWKTDSCQDSGGPL 380
QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431
RESULT 7
AAP60783
ID AAP60783 standard; Protein; 431 AA.
XX
XX AAP60783;
XX
XX 25-MAR-2003 (updated)
XX DT 23-OCT-1991 (first entry)
XX DE Human urokinase.
XX
XX E.coli; high molecular urokinase.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Protein 21..431
XX
XX JP61181377-A.
XX
XX 14-AUG-1986.
XX PD
XX 25-JAN-1985; 85JP-0011032.
XX PF
XX 25-JAN-1985; 85JP-0011032.
XX PR
XX (NISC) NISSAN CHEM IND LTD.
XX PA (HODO) HODOGAVA CHEM IND CO LTD.
XX PA (SAGA) SAGAMI CHEM RES CENTRE.
XX PA (CENG) CENTRAL GLASS CO LTD.
XX PA (NIPS) NIPPON SODA CO.
XX PA (TOYU) TOYO SODA MFG CO LTD.
XX
XX WPI: 1986-254744/39.
XX DR N-PSDB; AAN60703.
XX
XX Human urokinase gene - has N-end of aminoacid sequence coded by
XX PT codon used in Escherichia coli.
XX
XX Disclosure; Fig 2; 19pp; Japanese.
XX
XX The claimed gene product may be expressed in a transformed E.coli
XX CC host, for the efficient production of high molecular human urokinase.
XX CC The N-terminal of the protein expressed by the transforming plasmid
XX CC is replaced with a codon frequently used in E.coli.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX
XX SQ Sequence 431 AA;
Query Match 100.0%; Score 2301; DB 7; Length 431;
Best Local Similarity 100.0%; Pred. No. 5,7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYVQGLK 120
DB 81 KASDTMTGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRPRPCVYVQGLK 140
QY 121 PLVQECMVHDCADGKKPSPPEELKFCQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRRH 180
|||||

Db 141 PLVQECWVHDCADGKKPSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPWFAAIYRRH 200
 QY 181 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 240
 Db 201 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
 QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300
 Db 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 360
 Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 380
 QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
 Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 8

AAP70258
 ID AAP70258 standard; protein; 431 AA.

XX AAP70258;

AC 25-MAR-2003 (updated)

DT 19-MAY-1991 (first entry)

XX Sequence of human prourokinase and leader.

XX Cardiovascular disease treatment; fibrin affinity; thrombolytic;

KW enzyme; protease.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /label= leader

FT 21..431

FT /label= prourokinase

PN EP231883-A.

XX 12-AUG-1987.

XX 29-JAN-1987; 87EP-0101209.

XX 31-JAN-1986; 86JP-0017734.

PR 30-JAN-1987; 87JP-0018626.

XX (SAGA) SAGAMI CHEM RES CENTRE.

PA (NIPS) NIPPON SODA CO.

PA (CENG) CENTRAL GLASS CO LTD.

PA (TOXU) TOYO SODA MFG CO LTD.

PA (NISC) NISSAN CHEM IND LTD.

PA (NISC) NISSAN CHEMICAL INDS KK.

XX Tagawa M, Wada M, Yamada M, Yokoyama M, Numao M;

PI WPI; 1987-222882/32.

XX DR N-PSDB; AAN70390.

XX Hybrid plasminogen activator-like polypeptide - having a region

PT for affinity to fibrin from tissue plasminogen activator and a

PT region from prourokinase

XX Disclosure; Fig 2(1-5); 64pp; English.

XX The TPA portion of the claimed hybrid polypeptide (see FT) may

CC consist of 2 kringles from N-terminal first serine to 219th glycine

CC of human TPA, 1 kringle from 128th serine to 219th glycine of human

CC TPA or half a kringle from 161st methionine to 219th glycine (see

CC AAP70257). The C-terminal half of the hybrid polypeptide may contain
 CC an AA sequence from 150th glutamine to C-terminal 411th leucine of
 CC prourokinase (see AAP70258).
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 8; Length 431;

Best Local Similarity 100.0%; Pred. No. 5.7e-178;

Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHGVPSNCDCLNGCTCVSNKYFSNIHWNCNPKFGGQHCEIDKSKTCYEGNGHYRG 60

Db 21 SNELHGVPSNCDCLNGCTCVSNKYFSNIHWNCNPKFGGQHCEIDKSKTCYEGNGHYRG 80

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKENVYCRNPNRRRPPWCYVQVGLK 120

Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKENVYCRNPNRRRPPWCYVQVGLK 140

QY 121 PLVQECWVHDCADGKKPSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPWFAAIYRRH 180

Db 141 PLVQECWVHDCADGKKPSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPWFAAIYRRH 200

QY 181 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 240

Db 201 RGGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLI 260

QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 300

Db 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCIEITGFGK 320

QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 360

Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTKMLCAADPOWKTDSCQDGGPL 380

QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 9

AAP80430

ID AAP80430 standard; protein; 431 AA.

XX AAP80430;

XX 25-MAR-2003 (updated)

DT 14-SEP-1990 (first entry)

XX Deduced AA sequence of the single chain urokinase plasminogen activator

DE (SCU-PA) cDNA insert prepared from human Hep3 cells.

XX Single chain urokinase plasminogen activator (SCU-PA); human Hep3 cells;

KW glyceraldehyde-3-phosphate dehydrogenase gene; thrombosis prevention;

XX thrombosis treatment.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein 21..411

XX EP288435-A.

XX 26-OCT-1988.

XX 11-APR-1988; 88EP-0810234.

XX 15-APR-1987; 87GB-0009081.

PR 16-JUN-1987; 87GB-0014059.

PR 04-DEC-1987; 87IE-0003299.

XX (CIBA) CIBA GEIGY AG.

Db 261 LHKDYSADTLAHNDIALALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320
Qy 301 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 11
AAP92119
ID AAP92119 standard; protein; 431 AA.
XX AC AAP92119;
XX DT 25-MAR-2003 (updated)
XX DT 29-JUN-1990 (first entry)
XX DE Natural human prourokinase.
XX Human prourokinase; antithrombotic; derivative.
XX Homo sapiens.
XX FH Key Location/Qualifiers
FT Misc-difference 1 /note="Optional in new deriv."
FT Misc-difference 2..155 /note="Incorporated into new deriv."
FT Misc-difference 135 /note="May be replaced by a non-basic AA in new deriv."
FT Misc-difference 156 /note="Undefined residue in new deriv."
FT Misc-difference 157 /note="Pro, Gly, Ala or Val in new deriv."
FT Misc-difference 158 /note="Lys or Arg in new deriv."
XX W08901513-A.
XX 23-FEB-1989.
XX 18-AUG-1988; 88WO-JP00815.
XX 19-AUG-1987; 87JP-0204149.
XX (SAGA) SAGAMI CHEM RES CENTRE.
XX (CENG) CENTRAL GLASS CO LTD.
XX (HODO) HODOGAYA CHEM KK.
XX (NIPS) NIPPON SODA CO.
XX (NISC) NISSAN CHEM IND LTD.
XX Kobayashi Y, Omori M, Yamada C;
XX WPI; 1989-068869/09.
XX N-PSDB; AAN91075.
XX Antithrombotic fast-acting pro-urokinase deriv. -
XX produced by culture of E. coli transformant contg. new plasmid
XX of PMUT9Q family.
XX Disclosure; Fig 1; 75pp; Japanese.
XX A human prourokinase (PU) deriv. is new which is based upon residues
XX 2-155 of natural human prourokinase. The new deriv. is produced by
XX E. coli J103/PMUT9Q-RPK in culture. It is a fast-acting drug for
XX the treatment and prevention of thrombosis.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX (Updated on 25-MAR-2003 to correct PI field.)
XX Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHPYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHPYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNYCRPNDRRRPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGKPKSSPPEELKFCQGGOKTLRPFKIIIGGEFTTIENQPFMAIYRRH 180
Db 141 PLVQECMVHDCADGKPKSSPPEELKFCQGGOKTLRPFKIIIGGEFTTIENQPFMAIYRRH 200
Qy 181 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVLGRSLNSNTQGEKMEVENLI 240
Db 201 RGSVTVVCGSLISPCWVISATHCFIDYPKKEDIYVLGRSLNSNTQGEKMEVENLI 260
Qy 241 LHKDYSADTLAHNDIALALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 300
Db 261 LHKDYSADTLAHNDIALALKIRSEKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320
Qy 301 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVVKLLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 12
AAR04253
ID AAR04253 standard; protein; 431 AA.
XX AC AAR04253;
XX DT 25-MAR-2003 (updated)
XX DT 12-SEP-1990 (first entry)
XX Human pro-urokinase from the cDNA of clone pcUK176.
XX Non-glycosylated; pro-urokinase; E. coli; Ptip promoter; MS-2 RBS.
XX Synthetic.
XX EP365894-A.
XX 02-MAY-1990.
XX 06-OCT-1989; 89EP-0118586.
XX 11-OCT-1988; 88GB-0023833.
XX (FARM) FARMITALIA ERBA SPA CARLO.
XX Brandazza A, Sarmientos P, Orsini G;
XX WPI; 1990-133447/18.
XX N-PSDB; AAQ04107.
XX Non-glycosylated pro-urokinase prodn. - using E.coli B strains and E.coli
XX promoter Ptip and Shine-Dalgarno sequence MS-2.
XX Disclosure; Page ?; ?pp; English.
XX SER residue at position 21 is the start of the mature proUK.
XX Non-glycosylated proUK (MW 45kd) produced by E.coli B strain containing
XX the sequence..
XX See also AAQ04101-07.

CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 431 AA;
 SQ

Query Match 100.0%; Score 2301; DB 11; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5,7e-178;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140

QY 121 PLVQECMWHDCADGKKPSPPEELAFQCCQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
 DB 141 PLVQECMWHDCADGKKPSPPEELAFQCCQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200

QY 181 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
 DB 201 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 260

QY 241 LHKDYSADTLAHNDIALLKIRSKGRCACPQSRITQITCLPSMYNDPQGTSCITGFGK 300
 DB 261 LHKDYSADTLAHNDIALLKIRSKGRCACPQSRITQITCLPSMYNDPQGTSCITGFGK 320

QY 301 ENSTDYLYPEQIKMTVVKLIISHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 360
 DB 321 ENSTDYLYPEQIKMTVVKLIISHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 380

QY 361 VCSLQGRMTLTGIVSWGRGKALKDKPGVTVRVSHFLPWIRSHTKENGLAL 411
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVTVRVSHFLPWIRSHTKENGLAL 431

RESULT 14
 AAR07112
 ID AAR07112 standard; protein; 431 AA.
 AC AAR07112;
 XX
 XX
 DT 25-MAR-2003 (updated)
 DT 24-JAN-1991 (first entry)
 XX
 XX Human pro-Urokinase encoded by plasmid pUK1.
 DE pro-Urokinase; transgenic mice.
 KW Synthetic.
 OS EP390592-A.
 XX
 XX PD 03-OCT-1990.
 XX
 XX PF 30-MAR-1990; 90EP-0303445.
 XX
 XX PR 31-MAR-1989; 89JP-0078574.
 XX
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX (EXPE-) CENT INST EXPER ANIMALS.
 XX (JIKK-) JIKKEN DOBUTSU CHUO KENK.
 XX
 XX PI Sekine S, Ito S, Katsuki M;
 XX WPI; 1990-299492/40.
 DR N-ESDB; AAQ06049.
 DR
 XX Prodn. of recombinant protein, esp. human pro-urokinase - from
 PT milk of transgenic animals using promoter of bovine alpha S1 casein
 PT chromosomal gene.
 XX

Example; Table 1; 55pp; English.
 E.coli strain C600SF8 was transformed with recombinant plasmid
 containing ds DNA derived from human pharynx cancer cell strain
 Detroit 562. 10000 colonies were screened and one
 positive clone was identified. Plasmid pUK1 was isolated and found
 to contain the coding region and 3' non-coding region of pro-UK
 downstream of Cys(41). Four silent substitutions were identified
 c.f. Holmes et al., Biotechnology, vol.3, p.923 (1985) as follows:
 CC (sic) Asn(254), AAC to AAT; Leu(340), CTA to CTG; Pro(345), CCC to
 CC CCA; Gln(346), CAA to CAG.
 CC See also AAQ06045-Q06048 and AAQ06392.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 431 AA;
 SQ

Query Match 100.0%; Score 2301; DB 11; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5,7e-178;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
 DB 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140

QY 121 PLVQECMWHDCADGKKPSPPEELAFQCCQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 180
 DB 141 PLVQECMWHDCADGKKPSPPEELAFQCCQKTLRPRFKIIGGEFTTIENQPFAAIYRRH 200

QY 181 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 240
 DB 201 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFVENLI 260

QY 241 LHKDYSADTLAHNDIALLKIRSKGRCACPQSRITQITCLPSMYNDPQGTSCITGFGK 300
 DB 261 LHKDYSADTLAHNDIALLKIRSKGRCACPQSRITQITCLPSMYNDPQGTSCITGFGK 320

QY 301 ENSTDYLYPEQIKMTVVKLIISHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 360
 DB 321 ENSTDYLYPEQIKMTVVKLIISHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 380

QY 361 VCSLQGRMTLTGIVSWGRGKALKDKPGVTVRVSHFLPWIRSHTKENGLAL 411
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDKPGVTVRVSHFLPWIRSHTKENGLAL 431

RESULT 14
 AAR07903
 ID AAR07903 standard; Protein; 431 AA.
 AC AAR07903;
 XX
 XX
 DT 13-JUL-1994 (first entry)
 XX
 XX Pro-urokinase derivative.
 DE Pro-urokinase, half-life; thrombolytic; thrombosis; fibrinolytic;
 KW factor.
 XX
 XX OS Homo sapiens.
 XX
 XX FN JP05336965-A.
 XX
 XX PD 21-DEC-1993.
 XX
 XX PF 17-OCT-1991; 91JP-0269615.
 XX
 XX PR 17-OCT-1991; 91JP-0269615.
 XX
 XX PA (KYOW) KYOWA HAKKO KOGYO KK.

```
XX WPI; 1994-030907/04.
DR N-PSDB; AAQ55772.
XX
PT Novel human pro-urokinase derivs. having long half-life - with
PT high thrombolytic activity, useful for treatment of thrombosis
XX
PS Disclosure; Page 15-17; 29pp; Japanese.
XX
CC Sequences (AAQ55771-72) are pro-urokinase derivatives. The products
CC have an inserted sugar moiety having an amino acid substituted,
CC depleted or inserted variant around the thrombin cleavage site.
CC They also have a long half-life allowing them to be used in the
CC treatment of thrombosis.
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 2301; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMWHDCADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 180
DB 141 PLVQECMWHDCADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 181 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENMKFEVENLI 240
DB 201 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 300
DB 261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGGFL 360
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGGFL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWRSHTKBENGIAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTVRVSHFLPWRSHTKBENGIAL 431
```

RESULT 15
AAQ63141
ID AAQ63141 standard; Protein; 431 AA.

```
XX AAQ63141;
XX
DT 25-MAR-2003 (updated)
DT 09-JUN-1995 (first entry)
XX
DE Full length human urokinase protein.
XX
KW Human urokinase glycoproteins; cardiovascular diseases;
KW pulmonary embolism.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Sig peptide 1..20
FT Disulfide-bond 70..151
FT Disulfide-bond 91..133
FT Disulfide-bond 122..146
FT Disulfide-bond 168..299
FT Disulfide-bond 209..225
```

```
FT Disulfide-bond 217..288
FT Disulfide-bond 313..382
FT Disulfide-bond 345..361
FT Disulfide-bond 372..400
FT Cleavage-site 179..180
FT /note= "Cleavage of this site produces a bioactive
FT two chain form of urokinase"
```

XX EP620279-A1.

XX 19-OCT-1994.

XX 14-APR-1983; 94EP-0104777.

XX 15-APR-1982; 82US-0368773.

XX 14-MAR-1983; 83US-0474930.

XX 14-APR-1983; 83EP-0103629.

XX (GETH) GENENTECH INC.

XX Heyneker HL, Holmes WE, Vehar GA;

XX WPI; 1994-318362/40.

XX N-PSDB; AAQ73483.

XX Prodn. of human urokinase glycoproteins - using a recombinant
PT expression system used for the treatment of vascular diseases or
PT conditions.

XX Claim 1; Fig 4; 41pp; English.

XX AAQ73483 is the cDNA sequence which encodes AAR63141 the full length
CC 54000 dalton human urokinase (UK) protein. This cDNA was used in
CC the construction of a plasmid capable of transforming either yeast
CC or vertebrate cells, enabling them to produce the 54000 dalton
CC human UK protein. The UK glycoprotein produced could then be used
CC in the treatment of cardiovascular diseases, including pulmonary
CC embolism. The UK produced using this method had the advantage of a
CC specific activity towards fibrin and extant thrombi, not
CC demonstrated previously with UK isolated from natural sources.
CC (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 25-MAR-2003 to correct PF field.)
CC (Updated on 25-MAR-2003 to correct PR field.)

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMWHDCADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 180
DB 141 PLVQECMWHDCADGKKPSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPFWFAAIYRRH 200
QY 181 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENMKFEVENLI 240
DB 201 RGGSVTVYCGGSLISPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGENMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 300
DB 261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGGFL 360
DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGYGSEVTTKMLCAADPQWKTDSCQDGGGFL 380
```


QY 361 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWIRSHTKENGLAL 411
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 16

ID AAY99591 standard; protein; 431 AA.

XX AAY99591;

XX 13-SEP-2000 (first entry)

XX Human plasminogen activator urokinase, u-PA.

DE Human; serine protease; plasminogen activator; cardiant;

XX thrombolytic; heart attack; stroke; blood clotting disorder.

XX Homo sapiens.

XX WO200032759-A1.

XX 08-JUN-2000.

XX 06-MAY-1999; 99WO-US09991.

XX 02-DEC-1998; 98US-0110588.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX Lin X, Zhang XC, Tang JUN;

XX WPI; 2000-422975/36.

XX Polypeptide with plasminogen activator activity useful as thrombolytic agent for treating blood clot disorders e.g. heart attack, comprises 10 amino acid peptide fragment for recognition or activation of plasminogen -

XX Disclosure; Page 26-28; 41pp; English.

XX The present sequence is human plasminogen activator urokinase (u-PA), a serine protease which hydrolyses a peptide bond in human plasminogen to convert it to its active form, plasmin. Plasminogen is the principal serine protease zymogen in the extracellular fluids of vertebrates. Plasmin is implicated in pericellular proteolysis associated with a wide range of physiological and pathological processes. Plasminogen activators regulate plasminogen expression either by hydrolyzing a peptide bond, as in the case of u-PA, or by forming tight binding complexes with plasminogen to spontaneously convert it to plasmin. Review of sequence homologies of several plasminogen activators and chymotrypsin has identified a six amino acid peptide involved in plasminogen activation. This peptide is particularly useful when inserted between amino acid residues 644 and 645 of full length human plasminogen. Novel plasminogen activators have been made based upon the plasminogen activation/recognition site of plasminogen binding proteins. The polypeptides are useful in preparing thrombolytic agents for treating blood clotting disorders such as heart attack.

XX Sequence 431 AA;

XX Query Match 100.0%; Score 2301; DB 21; Length 431;
 XX Best Local Similarity 100.0%; Pred. No. 5,7e-178;
 XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHLQVPSNCDLNGCTCVSNKYFSNHWNCNPKFGGQHCEDKSKTCYEGNGHYRG 60
 DB 21 SNEHLQVPSNCDLNGCTCVSNKYFSNHWNCNPKFGGQHCEDKSKTCYEGNGHYRG 80

QY 61 KASDTDMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 81 KASDTDMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPNRRPWCYVQVGLK 140

QY 121 PLVQECMWVHDCADGKXPSPPPEELKFCQGOKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 180
 DB 141 PLVQECMWVHDCADGKXPSPPPEELKFCQGOKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 200
 QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVLGSRSLNSNTQGMKEVENLI 240
 DB 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDIYVLGSRSLNSNTQGMKEVENLI 260
 QY 241 LHKDYSADTLAHHNDIALLLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 300
 DB 261 LHKDYSADTLAHHNDIALLLKIRSKGRCAQPSRTIQTICLPSMYNDPQFGTSCBITGFGK 320
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGPL 360
 DB 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKTDSCQSDSGPL 380
 QY 361 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWIRSHTKENGLAL 411
 DB 381 VCSLQGRMTLTGIVSWGRGKALKDPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 17

AAV50869 standard; protein; 431 AA.

XX AAY50869;

XX 24-FEB-2000 (first entry)

XX Human urokinase protein fragment.

XX Urokinase; human; thrombolytic agent; streptokinase; antigenic;

XX blood clot; heart attack; treatment.

XX Homo sapiens.

XX WO9957251-A2.

XX 11-NOV-1999.

XX 06-MAY-1999; 99WO-US10086.

XX 06-MAY-1998; 98US-0084392.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX Zhang XC, Lin X, Tang JUN;

XX WPI; 2000-052966/04.

XX New thrombolytic agents derived from modified humanized streptokinase, useful for treating blood clot disorders -

XX Disclosure; Page 46-48; 55pp; English.

XX This invention describes a novel thrombolytic agent comprising streptokinase where at least one nonessential portion has been modified. The invention also describes a method of forming a thrombolytic agent comprising determining a nonessential portion of streptokinase and modifying the nonessential portion to render the resulting protein less antigenic. The modified streptokinase is used to treat blood clot disorders, such as heart attacks. The modified streptokinase has less antigenicity than streptokinase but is still able to complex plasminogen and lead to plasminogen activation. Modified streptokinase with the nonessential portions removed or truncated simplify the molecule. Such smaller proteins are cheaper and easier to produce. This sequence represents a fragment of the human urokinase protein which is used in the description of the method of the invention.

XX Sequence 431 AA;

XX Query Match 100.0%; Score 2301; DB 21; Length 431;

Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTCYEGNGHYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWSNATVLQOTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSNATVLQOTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECWHDCAKPKSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPFWFAAIYRRH 180
DB 141 PLVQECWHDCAKPKSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPFWFAAIYRRH 200
QY 181 RGSVTVYVCGSLISPCWISATHCDFIDYPKEDYIVYLSRSLNSNTQEMKFEVENLI 240
DB 201 RGSVTVYVCGSLISPCWISATHCDFIDYPKEDYIVYLSRSLNSNTQEMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIOTICLPSMYNDPQFTSCIEITGFGK 300
DB 261 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIOTICLPSMYNDPQFTSCIEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 18
AAB84605
ID AAB84605 standard; Protein; 431 AA.
XX AAB84605;
XX
XX 05-SEP-2001 (first entry)
XX
DE Amino acid sequence of urokinase plasminogen activator.
XX
KW Growth factor; protein inhibitor; protease; damaged tissue;
KW platelet-derived growth factor; PDGF; fibroblast growth factor; RGF;
KW connective tissue derived growth factor; CTGF; chrysalin; VEGF;
KW keratinocyte-derived growth factor; KGF; epidermal growth factor; EGF;
KW transforming growth factor-beta; TGF-beta; matrix metalloproteinase; MMP;
KW granulocyte macrophage colony stimulating factor; GM-CSF; uPA;
KW vascular endothelial growth factor; urokinase plasminogen activator;
KW dermal ulcer; wound.
XX
XX Homo sapiens.
OS
XX
XX WO200149309-A2.
XX
XX 12-JUL-2001.
XX
XX 21-DEC-2000; 2000WO-1B01935.
XX
XX 29-DEC-1999; 99GB-0030768.
XX
XX (PFIZ) PFIZER LTD.
PA (PFIZ) PFIZER INC.
XX
XX Davies MJ, Huggins JP, McIntosh PS, Occleston NL;
PI
XX WPI; 2001-418351/44.
DR
XX N-PSDB; AAH28220.
DR
XX Composition for the treatment of damaged tissue i.e. chronic wounds and
PT dermal ulcers comprises an inhibitor agent i.e. a protease and a growth
PT factor -
XX

PS Disclosure; Page 550; 572pp; English.
XX
CC The specification describes a pharmaceutical composition, comprising
CC a growth factor, an inhibitor agent, i.e. a protease. The inhibitor
CC agent inhibits the action of at least one specific adverse protein,
CC i.e. a protease, that is upregulated in a damaged tissue such as a
CC wound environment. Growth factors which are included in the composition
CC of the invention are platelet-derived growth factor (PDGF), fibroblast
CC growth factor (FGF), connective tissue derived growth factor (CTGF),
CC keratinocyte-derived growth factor (KGF), transforming growth
CC factor-beta (TGF-beta), granulocyte macrophage colony stimulating factor
CC (GM-CSF), epidermal growth factor (EGF), vascular endothelial growth
CC factor (VEGF), and chrysalin. Inhibitors which are included in the
CC composition of the invention include inhibitors of urokinase-type
CC plasminogen activator (uPA) and matrix metalloproteinase (MMP). The
CC composition is useful for the treatment of chronic damaged tissue, i.e.
CC wounds and dermal ulcers. The present sequence represents a human uPA,
CC and is used to produce the composition of the invention.
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 2301; DB 22; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTCYEGNGHYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWSNATVLQOTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 120
DB 81 KASTDTMGRPCLPWSNATVLQOTYHAHRSALQGLGKHNCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECWHDCAKPKSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPFWFAAIYRRH 180
DB 141 PLVQECWHDCAKPKSSPPELKFQCGQKTLRPRFKIIGGFTTIENQPFWFAAIYRRH 200
QY 181 RGSVTVYVCGSLISPCWISATHCDFIDYPKEDYIVYLSRSLNSNTQEMKFEVENLI 240
DB 201 RGSVTVYVCGSLISPCWISATHCDFIDYPKEDYIVYLSRSLNSNTQEMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIOTICLPSMYNDPQFTSCIEITGFGK 300
DB 261 LHKDYSADTLAHNDIALLKIRSGRCAQPSRTIOTICLPSMYNDPQFTSCIEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDYLYPEQLKMTVVKLIHRECCQPHYVGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 19
AAG79460
ID AAG79460 standard; Protein; 431 AA.
XX AAG79460;
XX
XX 15-NOV-2002 (first entry)
DT
XX
DE sc-uPA.
XX
KW Single chain prepro-urokinase; sc-uPA; pro-urokinase; HIV;
KW high molecular weight urokinase-type plasminogen activator; HMW-uPA;
KW long A; long B; EGF-like domain; kringle domain; urokinase receptor;
KW low molecular weight urokinase-type plasminogen activator; LMW-uPA;
KW CD87; binding domain.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers

FT Peptide 1..20
 FT /note= "Signal peptide"
 FT Protein 21..431
 FT /label= "Pro-urokinase/HMW-uPA"
 FT /note= "With a cleavage between amino acids 178-179"
 FT Protein 156..431
 FT /label= "LMW-uPA"
 FT /note= "With a cleavage between amino acids 178-179"
 XX EPI232755-A2.
 XX
 XX 21-AUG-2002.
 XX
 XX 15-FEB-2002; 2002EP-0003555.
 XX
 XX 20-FEB-2001; 2001JP-0042655.
 XX
 XX 19-JUN-2001; 2001JP-0184284.
 XX
 XX (JCRP-) JCR PHARM CO LTD.
 XX
 XX Wada M, Wada N;
 XX
 XX WPI; 2002-610512/66.
 XX
 XX N-PSDB; ABA00207.
 XX
 XX Anti-HIV agents, comprises ligand molecule that binds to CD87, e.g.
 PT high molecular weight urokinase-type plasminogen activator,
 PT amino-terminal fragment or an anti-CD87 antibody -
 XX
 XX Disclosure; Page 20-23; 38pp; English.
 XX
 XX This sequence represents single chain prepro-urokinase (sc-uPA).
 CC Pro-urokinase (amino acids 21-431) with a cleavage between amino
 CC acids 178 and 179 gives high molecular weight urokinase-type
 CC plasminogen activator (HMW-uPA). HMW-uPA is a protein consisting
 CC of two peptide chains linked by a di-sulphide bond. The chains,
 CC long A and B, are formed by enzymatic cleavage between amino acids
 CC 178 and 179 of pro-urokinase. HMW-uPA includes an RGF-like domain,
 CC a kringle domain and a urokinase receptor (CD87) binding domain.
 CC HMW-uPA is then cleaved between amino acids 155 and 156 to give low
 CC molecular weight urokinase-type plasminogen activator (LMW-uPA)
 CC (amino acids 156-178 and 179-431), that has no plasminogen activator
 CC activity. sc-uPA, or fragments of it, may be used in the anti-HIV
 CC agents of the invention which comprise a ligand molecule that binds to
 CC CD87. The agents are useful for treating HIV-infected humans for
 CC suppression of reproduction of HIV. The anti-HIV agents act by a
 CC mechanism of action different from those of conventional drugs, of
 CC widening the choice of therapeutics agents and avoiding problems of
 CC resistant HIV.
 XX
 XX Sequence 431 AA;
 SQ
 Query Match 100.0%; Score 2301; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCHIDKSKTCYEGNGHYRG 60
 DB 21 SNELHQVPSNCDLNGTGVSNKYFSNIHWCNCPKFGGQHCHIDKSKTCYEGNGHYRG 80
 QY 61 KASTDTWGRCLPWNSTAVLQOYTHAHSALQGLGKHNYCNPNRRPWCYVQVGLK 120
 DB 81 KASTDTWGRCLPWNSTAVLQOYTHAHSALQGLGKHNYCNPNRRPWCYVQVGLK 140
 QY 121 PLVCEMVDHCDGKPSPEELKFGGQKTLRPFKIIIGGEFTTIENQPFWAAIYRRH 180
 DB 141 PLVCEMVDHCDGKPSPEELKFGGQKTLRPFKIIIGGEFTTIENQPFWAAIYRRH 200
 QY 181 RGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQEMKFEVENLI 240
 DB 201 RGSVTVYCGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQEMKFEVENLI 260
 QY 241 LHKDYSADTLAHHNDIALKIRSEGRCAQPSRTIQICLPSMYNDPQFGTSCITGFGK 300

Db 261 LHKDYSADTLAHHNDIALKIRSEGRCAQPSRTIQICLPSMYNDPQFGTSCITGFGK 320
 QY 301 ENSTDYLYPEQLKMTVVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSDGGL 360
 Db 321 ENSTDYLYPEQLKMTVVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDSDGGL 380
 QY 361 VCSLQGRMTLGIIVSWGRGCALCKDKPGVYTRVSHFLPWIRSHRTEENGLAL 411
 Db 381 VCSLQGRMTLGIIVSWGRGCALCKDKPGVYTRVSHFLPWIRSHRTEENGLAL 431
 RESULT 20
 AAU99228
 ID AAU99228 standard; Protein; 431 AA.
 XX
 AC AAU99228;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 XX Human plasminogen activator, urokinase (PLAU).
 DE Human; Plasminogen activator; urokinase; PLAU; cancer; enzyme;
 KW cytostatic; serine protease; thrombolytic disorder; isogene;
 KW pulmonary embolism; chromosome 10q24-qter; haplotype; genotype;
 KW SNP; single nucleotide polymorphism; thrombolytic; gene therapy.
 XX
 OS Homo sapiens.
 XX
 XX WO200240503-A2.
 XX
 XX 23-MAY-2002.
 XX
 XX 14-NOV-2001; 2001WO-US44001.
 XX
 XX 17-NOV-2000; 2000US-249703P.
 XX
 XX (GENA-) GENAISSANCE PHARM INC.
 PI Anastasio AE, Bentivegna SC, Koshy B;
 XX
 XX WPI; 2002-519370/55.
 DR N-PSDB; ABK86597, ABK86598.
 XX
 XX Genetic variants of Plasminogen activator, Urokinase (PLAU) isogenes,
 PT useful for improving efficiency and reliability in drug development for
 PT treating thrombolytic disorders and cancer -
 XX
 XX Claim 27; Fig 3; 92pp; English.
 XX
 XX The invention relates to a polynucleotide comprising a first nucleotide
 CC sequence (NSI) comprising a PLAU (plasminogen activator, urokinase,
 CC a serine protease) isogene selected from isogenes 1-9 and 11-20 given
 CC in the specification, where each isogene comprises the regions of the
 CC PLAU gene or cDNA and is further defined by the corresponding sequence of
 CC polymorphisms (defining single nucleotide polymorphisms, SNP). Also
 CC included are methods of haplotyping/genotyping (and predicting the
 CC haplotype/genotype of the PLAU gene of an individual, identifying an
 CC association between a trait and at least one haplotype or haplotype pair
 CC of the PLAU gene, an isolated oligonucleotide for detecting a
 CC polymorphism in the PLAU gene, a recombinant non-human organism
 CC transformed or transfected with the gene or cDNA, fragments of the
 CC polynucleotides of at least 10 base pairs encompassing a polymorphic
 CC site, an isolated polymorphic variant PLAU protein or fragment, an
 CC isolated monoclonal antibody specific for PLAU, a computer system for
 CC storing and analysing polymorphism data for the PLAU gene and a genome
 CC anthology for the PLAU gene. PLAU is useful in screening for drugs and
 CC targeting PLAU that are useful for treating thrombolytic disorders and
 CC cancers. The methods are useful for improving the efficiency and
 CC reliability of the discovery and development of drugs for treating
 CC diseases associated with PLAU activity, in validating PLAU as a drug
 CC target and in the design of clinical trials for treating a specific
 CC condition of disease associated with PLAU activity. The antibody is

CC useful in diagnostic, prognostic and therapeutic methods. PLAU
CC polynucleotides are useful in studying the expression and function of
CC PLAU, and in expressing PLAU protein for use in screening for candidate
CC drugs to treat diseases related to PLAU activity. The gene for PLAU
CC is located on chromosome 10q24-qter. The present sequence represents the
CC PLAU protein.
XX
XX

SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 23; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRPNDRRPWCYVQVGLK 120
DB 81 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRPNDRRPWCYVQVGLK 140
QY 121 PLVQECWVHDCADGKSPPELKFQCGQKTLRPRFKIIGGFTTIENTOPWFAAIYRRH 180
DB 141 PLVQECWVHDCADGKSPPELKFQCGQKTLRPRFKIIGGFTTIENTOPWFAAIYRRH 200
QY 181 RGGSVTVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 240
DB 201 RGGSVTVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 300
DB 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 360
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 21

AAE17128
ID AAE17128 standard; Protein; 431 AA.

XX
XX
AC AAE17128;

DT 18-APR-2002 (first entry)

XX
DE Human uPA protein.

XX Human; cancer; urokinase-type plasminogen activator; uPA; inflammation;
KW Bts-1 transcription factor; N-acetylglucosaminyltransferase V; Gnt-V;
KW matrix-type metalloproteinase; MMP-1; MMP-3; gene therapy.
XX

OS Homo sapiens.

XX
XX WO200196606-A2.

XX
XX 20-DEC-2001.

XX
XX 14-JUN-2001; 2001WO-US19248.

XX
XX 14-JUN-2000; 2000US-0593488.

XX (NYXI-) NYXIS NEURO THERAPIES INC.

XX Yamamoto H, Kroes R, Moskal JR;

XX WPI; 2002-130746/17.

DR N-PSDB; AAD27855.

XX

PT Identifying a compound for treating cancer, comprises detecting
PT transcription factor Bts-1, N-acetylglucosaminyltransferase V,
PT urokinase-type plasminogen activator, matrix-type metalloproteinase-1
PT and -3 gene expression -
XX
XX

PS Example 1; Page 62-63; 63pp; English.

XX The invention relates to a method of identifying a compound for treating
CC cancer. The method involves detecting the expression of a panel of
CC sequences selected from transcription factor Bts-1, urokinase-type
CC plasminogen activator (uPA), N-acetylglucosaminyltransferase V (Gnt-V),
CC matrix-type metalloproteinase (MMP)-1 and MMP-3 in the cell. The method
CC is useful for identifying a compound that affects a cell, particularly a
CC cancer cell or glioma cell, or a cell that is involved in inflammation.
CC It is used for diagnosis and/or treating cancer or other conditions that
CC are affected by one or more members of a panel of genes or their protein
CC product. The method is also useful for drug discovery, drug safety
CC evaluations and in gene therapy. The present sequence is human uPA
CC protein.
XX

SQ Sequence 431 AA;

Query Match 100.0%; Score 2301; DB 23; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.7e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRPNDRRPWCYVQVGLK 120
DB 81 KASTDTMGRCPLPWSATVLQOTYHAHRSDALQGLGKHNYCRPNDRRPWCYVQVGLK 140
QY 121 PLVQECWVHDCADGKSPPELKFQCGQKTLRPRFKIIGGFTTIENTOPWFAAIYRRH 180
DB 141 PLVQECWVHDCADGKSPPELKFQCGQKTLRPRFKIIGGFTTIENTOPWFAAIYRRH 200
QY 181 RGGSVTVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 240
DB 201 RGGSVTVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 300
DB 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQFGTSCITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 360
DB 321 ENSTDYLYPEQLKMTVVKLISHRECOQPHYVGSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

RESULT 22

ABU56547

ID ABU56547 standard; Protein; 431 AA.

XX
XX
AC ABU56547;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polypeptide #140.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX

OS Unidentified.

XX

PN W0200286443-A2.
 XX 31-OCT-2002.
 XX 18-APR-2002; 2002WO-US12476.
 XX 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Aziz N, Murray R;
 XX WPI; 2003-093161/08.
 DR N-PSDB; ABX76275.
 XX
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer
 XX
 XX Claim 27; Page 296; 453pp; English.
 XX
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.
 XX
 XX Sequence 431 AA;
 XX Query Match 100.0%; Score 2301; DB 24; Length 431;
 XX Best Local Similarity 100.0%; Pred. No. 5.7e-178;
 XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDLNGTGVSNKVFNSIHWNCNPKFGGQHCEDKSKTCYEGNGHYRG 60
 DB 21 SNEHQVPSNCDLNGTGVSNKVFNSIHWNCNPKFGGQHCEDKSKTCYEGNGHYRG 80
 QY 61 KASTDTMGRPCLPNNSATVLOQTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
 DB 81 KASTDTMGRPCLPNNSATVLOQTHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 140
 QY 121 PLVQCMVHDCADGKPPSPPELKFQCGQKTLPRFKIIGGEFTTIENQWFAAIYRRH 180
 DB 141 PLVQCMVHDCADGKPPSPPELKFQCGQKTLPRFKIIGGEFTTIENQWFAAIYRRH 200
 QY 181 RGGSVTYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKEVENLI 240
 DB 201 RGGSVTYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTQGMKEVENLI 260
 QY 241 LHKDYSADTLAHNDIALKIRSKRCAQPSRTIQTICLPSMWNDPFGTSCBITGFGK 300
 DB 261 LHKDYSADTLAHNDIALKIRSKRCAQPSRTIQTICLPSMWNDPFGTSCBITGFGK 320

QY 301 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYGVSEVTTMCLCAADPQWKTDSCQDSGGPL 360
 DB 321 ENSTDYLYPEQLKMTVVVKLISHRECCQPHYGVSEVTTMCLCAADPQWKTDSCQDSGGPL 380
 QY 361 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKENGAL 411
 DB 381 VCSLQGRMTLTGIVSWGRGCALDKPGVYTRVSHFLPWIRSHTKENGAL 431
 RESULT 23
 ABUS6708
 ID ABUS6708 standard; Protein; 431 AA.
 XX
 AC ABUS6708;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #301.
 XX
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 PN W0200286443-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 18-APR-2002; 2002WO-US12476.
 XX
 PR 18-APR-2001; 2001US-284770P.
 PR 10-MAY-2001; 2001US-290492P.
 PR 09-NOV-2001; 2001US-339245P.
 PR 13-NOV-2001; 2001US-350666P.
 PR 29-NOV-2001; 2001US-334370P.
 PR 12-APR-2002; 2002US-372246P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX Aziz N, Murray R;
 XX WPI; 2003-093161/08.
 DR N-PSDB; ABX76437.
 XX
 XX Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer
 XX
 PS Claim 27; Page 424; 453pp; English.
 XX
 CC The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung
 CC cancer-associated polynucleotides and polypeptides are used for
 CC identifying a compound that modulates a lung cancer-associated
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated
 CC cell to treat lung cancer in a patient and for treating a mammal having
 CC lung cancer by administering a modulatory compound identified. The
 CC methods are useful for treating lung cancer, such as small cell lung
 CC cancer, non-small cell lung cancer or other benign or precancerous
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides
 CC and polypeptides are useful for diagnostic purposes and as targets for
 CC screening for therapeutic compounds that modulate lung cancer, such as
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated
 CC polypeptides of the invention.

```
XX Sequence 431 AA;
XX Query Match 100.0%; Score 2301; DB 24; Length 431;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-178;
XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLIQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLIQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEFTTTIENOPWFAAIYRRH 180
DB 141 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEFTTTIENOPWFAAIYRRH 200
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLGSRSLNSNTQGMKFEVENLI 240
DB 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLGSRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 300
DB 261 LHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVVKLISHRECQOPHYVYGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDYLYPEQLKMTVVVKLISHRECQOPHYVYGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 24
ABU11076
ID ABU11076 standard; Protein; 431 AA.
AC ABU11076;
XX
XX 05-FEB-2003 (first entry)
XX Human urokinase plasminogen activator.
XX
XX Urokinase plasminogen activator; gene therapy; cancer;
XX hyperproliferative disorder; cancer; breast cancer; colon cancer;
XX bone cancer; brain cancer; ovary cancer; cervix cancer;
XX endometrium cancer; stomach cancer; kidney cancer; tumour metastasis.
XX
XX Homo sapiens.
XX
XX WO200279515-A1.
XX
XX 10-OCT-2002.
XX
XX 18-MAR-2002; 2002WO-US08112.
XX
XX 30-MAR-2001; 2001US-0821972.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Baker BF, Freier SM, Watt AT;
XX
XX WPI: 2003-058441/05.
XX N-PSDB; ABX17681.
XX
XX New antisense compound, useful for preparing a composition for treating
XX hyperproliferative disorders, cancer e.g., breast, colon, bone, brain,
XX ovary, cervix, endometrium, stomach or kidney cancer, or tumor
XX metastasis -
XX
```

```
PS Disclosure; Page 101-102; 153pp; English.
XX
XX A new compound, which is 8-50 nucleobases in length targeted
XX to a nucleic acid molecule encoding urokinase plasminogen activator,
XX specifically hybridises with and inhibits the expression of urokinase
XX plasminogen activator. The compound is useful for preparing a
XX composition for treating (e.g. by gene therapy) hyperproliferative
XX disorder, cancer e.g., breast, colon, bone, brain, ovary, cervix,
XX endometrium, stomach or kidney cancer, or tumour metastasis. This
XX is the amino acid sequence of a urokinase plasminogen activator.
XX Sequence 431 AA;
XX Query Match 100.0%; Score 2301; DB 24; Length 431;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-178;
XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLIQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLIQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEFTTTIENOPWFAAIYRRH 180
DB 141 PLVQECMVHDCADGKPKSSPPEELKFQCGQKTLRPRFKIIGGEFTTTIENOPWFAAIYRRH 200
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLGSRSLNSNTQGMKFEVENLI 240
DB 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVVLGSRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 300
DB 261 LHKDYSADTLAHHNDIALLKIRSKEGRCAPSRITQITCLPSMYNDPQFGTSCBITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVVKLISHRECQOPHYVYGSEVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDYLYPEQLKMTVVVKLISHRECQOPHYVYGSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411
DB 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 25
AAR20537
ID AAR20537 standard; Protein; 434 AA.
XX
XX AAR20537;
XX
XX 25-MAR-2003 (updated)
XX 21-MAY-1992 (first entry)
XX
XX Amidated deriv. of pro-urokinase (2).
XX
XX Pro-urokinase; plasminogen activator; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein 1..431
XX /label= pro-UK
XX
XX DE4122688-A.
XX
XX 16-JAN-1992.
XX
XX 09-JUL-1991; 91DE-4122688.
XX
XX 12-JUL-1990; 90GB-0015369.
XX 10-JUL-1991; 91GB-0014846.
XX
```

XX (FARM) FARMITALIA ERBA SRL CARLO.
 XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
 XX WPI; 1992-025815/04.
 XX New amidated derivs. of human pro-urokinase - are fibrinolytic
 PT and can be used to treat acute myocardial infarction, pulmonary
 PT embolism or deep venous thrombosis
 XX Claim 4,8; Page 8; 18pp; German.
 XX The protein has fibrinolytic activity and can be used in the same
 CC way as PUK, e.g. for treating acute myocardial infarction, lung
 CC embolism and deep venous thrombosis. It has greater affinity for
 CC plasminogen bound to fibrin than for circulating plasminogen, so
 CC have high selectivity for thrombi with reduced chance of bleeding.
 CC Compared with the COOH-terminated cpds., it has a better stability
 CC against most carboxypeptidases and prolonged half life.
 CC See also AAQ20360, AAQ20754-58 and AAR20536-38.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX Sequence 434 AA;
 XX Query Match 100.0%; Score 2301; DB 13; Length 434;
 XX Best Local Similarity 100.0%; Pred. No. 5.7e-178;
 XX Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECWHDCAADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 180
 Db 141 PLVQECWHDCAADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
 QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 240
 Db 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 260
 QY 241 LHKDYSADTLAHHNDIALIKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 300
 Db 261 LHKDYSADTLAHHNDIALIKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 320
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPL 360
 Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPL 380
 QY 361 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGIAL 411
 Db 381 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGIAL 431
 RESULT 26
 AAR20538
 ID AAR20538 standard; Protein; 434 AA.
 XX AAR20538;
 XX AC
 XX 25-MAR-2003 (updated)
 DT 21-MAY-1992 (first entry)
 XX Amidated deriv. of pro-urokinase (3).
 XX Pro-urokinase; plasminogen activator; ss.
 XX Homo sapiens.

Key Location/Qualifiers
 Protein 1..431
 /label= pro-UK
 DB4122688-A.
 16-JAN-1992.
 09-JUL-1991; 91DE-4122688.
 12-JUL-1990; 90GB-0015369.
 10-JUL-1991; 91GB-0014846.
 (FARM) FARMITALIA ERBA SRL CARLO.
 Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
 WPI; 1992-025815/04.
 New amidated derivs. of human pro-urokinase - are fibrinolytic
 and can be used to treat acute myocardial infarction, pulmonary
 embolism or deep venous thrombosis
 Claim 4,9; Page 8; 18pp; German.
 The protein has fibrinolytic activity and can be used in the same
 way as PUK, e.g. for treating acute myocardial infarction, lung
 embolism and deep venous thrombosis. It has greater affinity for
 plasminogen bound to fibrin than for circulating plasminogen, so
 have high selectivity for thrombi with reduced chance of bleeding.
 Compared with the COOH-terminated cpds., it has a better stability
 against most carboxypeptidases and prolonged half life.
 See also AAQ20360, AAQ20754-58 and AAR20536-38.
 (Updated on 25-MAR-2003 to correct PA field.)
 Sequence 434 AA;
 Query Match 100.0%; Score 2301; DB 13; Length 434;
 Best Local Similarity 100.0%; Pred. No. 5.7e-178;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
 Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 Db 81 KASTDTMGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
 QY 121 PLVQECWHDCAADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 180
 Db 141 PLVQECWHDCAADGKKSPPEELKFCGQKTLRPRFKIIGGEFTTIENQPMFAAIYRRH 200
 QY 181 RGGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 240
 Db 201 RGGSVTVVCGGSLISPCWVISATHCFIDYPKEDYIVLGRSRLNSNTQGMKFEVENLI 260
 QY 241 LHKDYSADTLAHHNDIALIKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 300
 Db 261 LHKDYSADTLAHHNDIALIKIRSKGRCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 320
 QY 301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPL 360
 Db 321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYGGSEVTTKMLCAADPQWKTDSCQDGGPL 380
 QY 361 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGIAL 411
 Db 381 VCSLQGRMTLTGIVSWGRGCAKDKPGVYTRVSHFLPWIRSHTKKEENGIAL 431
 RESULT 27
 AAR20536
 ID AAR20536 standard; Protein; 436 AA.

XX AC AAR20536;
XX 25-MAR-2003 (updated)
XX 21-MAY-1992 (first entry)
XX Amidated deriv. of pro-urokinase (1).
XX Pro-urokinase; plasminogen activator; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Protein 1..431
XX /label= pro-UK
XX /note= "the amidated form of pro-UK is
XX /note= "claimed in claim 1"
XX Misc-difference 433..436
XX /note= "may be any amino acid, pref. Lys or Arg,
XX or 0-4 amino acids may be omitted"
XX DE4122688-A.
XX 16-JAN-1992.
XX 09-JUL-1991; 91DE-4122688.
XX 12-JUL-1990; 90GB-0015369.
XX 10-JUL-1991; 91GB-0014846.
XX (FARM) FARMITALIA ERBA SRL CARLO.
XX Gozzini L, Visco C, Perego R, Roncucci R, Sarmientos P;
XX WPI; 1992-025815/04.
XX New amidated derivs. of human pro-urokinase - are fibrinolytic
XX and can be used to treat acute myocardial infarction, pulmonary
XX embolism or deep venous thrombosis
XX Claim 1,4,7; Page 8; 18pp; German.
XX The protein has fibrinolytic activity and can be used in the same
XX way as PUK, e.g. for treating acute myocardial infarction, lung
XX embolism and deep venous thrombosis. It has greater affinity for
XX plasminogen bound to fibrin than for circulating plasminogen, so
XX have high selectivity for thrombi with reduced chance of bleeding.
XX Compared with the COOH-terminated cpds., it has a better stability
XX against most carboxypeptidases and prolonged half life.
XX See also AAO20360, AAO20754-58 and AAR20536-38.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 436 AA;
Query Match 100.0%; Score 2301; DB 13; Length 436;
Best Local Similarity 100.0%; Pred. No. 5.8e-178;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHVPNCCLNGTCTVSKNYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHYRG 60
DB 21 SNELHVPNCCLNGTCTVSKNYFSNIHWCNPKKFGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTWGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
DB 81 KASTDTWGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPNRRRPMCYVQVGLK 140
QY 121 PLVQECWHDCAKGPSPPELKFQCGQKTLRPFKLIIGGEFTTIENQPFWFAIYRRH 180
DB 141 PLVQECWHDCAKGPSPPELKFQCGQKTLRPFKLIIGGEFTTIENQPFWFAIYRRH 200
QY 181 RGSSTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLI 240
DB 201 RGSSTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFEVENLI 260

QY 241 LHKDYSADTLAHHNDIALLKIRSKRCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 300
DB 261 LHKDYSADTLAHHNDIALLKIRSKRCAQPSRTIQTICLPSMYNDPQGTSCITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPL 360
DB 321 ENSTDYLYPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPQWKTDSCQDGGPL 380
QY 361 VCSLOGRMTLTGIVSGRGCAKDKPGVYTVRVSHPLPWIRSHRTKEENGIAL 411
DB 381 VCSLOGRMTLTGIVSGRGCAKDKPGVYTVRVSHPLPWIRSHRTKEENGIAL 431
RESULT 28
AAW24578
ID AAW24578 standard; Protein; 430 AA.
XX AAW24578;
XX AC AAW24578;
XX 25-MAR-2003 (updated)
XX 11-NOV-1997 (first entry)
XX Inhibitor resistant urokinase.
XX Urokinase; inhibitor resistant; plasminogen; human; whey acid protein;
XX plasminogen activator inhibitor-1; plasminogen activator; protease; WAP;
XX plasmin; proteolytic enzyme; clot lysis; heart attack; heart muscle.
XX Homo sapiens.
XX Synthetic.
XX Key Location/Qualifiers
XX Peptide 1..19
XX /note= "whey acid protein signal peptide"
XX Protein 20..430
XX /note= "urokinase"
XX Domain 20..64
XX /note= "B-domain"
XX Domain 61..150
XX /note= "Kringle-1 domain"
XX Domain 179..430
XX /note= "P-domain"
XX Misc-difference 198..203
XX /note= "deleted in modified urokinase of the invention"
XX US648253-A.
XX 15-JUL-1997.
XX 08-SEP-1992; 92US-0942157.
XX 20-DEC-1990; 90US-0631673.
XX 08-SEP-1992; 92US-0942157.
XX (TSIT-) TSI CORP.
XX Wei C;
XX WPI; 1997-372062/34.
XX N-PSDB; AAT80075.
XX Deletion-modified urokinase protein - with increased resistance to
XX inhibition by plasminogen activator inhibitor-1
XX Disclosure; Column 15-18; 16pp; English.
XX This sequence represents the full length urokinase, including the
XX whey acid protein (WAP) signal peptide. This sequence has residues
XX 179-184 of the urokinase sequence deleted to create the modified
XX urokinase of the invention. The modified urokinase (see AAW24579)
XX cleaves plasminogen, and has a lower binding affinity for plasminogen
XX activator inhibitor-1 than the corresponding unmodified urokinase.

CC Urokinase is one of two types of mammalian plasminogen activators (PA),
 CC the other being tissue type PA. PAs catalyse the conversion of the
 CC circulating zymogen plasminogen to the broad spectrum protease plasmin by
 CC limited proteolysis. The modified urokinase can be used for clot lysis,
 CC specifically to dissolve heart attack-causing clots before they cause
 CC permanent damage to heart muscle. The urokinase mutant is more resistant
 CC to inhibition by plasminogen activator inhibitors than the unmodified
 CC urokinase. It can be selectively expressed and secreted from the mammary
 CC glands of transgenic animals.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX Sequence 430 AA;

Query Match 99.9%; Score 2298; DB 18; Length 430;
 Best Local Similarity 99.8%; Pred. No. 9.9e-178;
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60
 DB 20 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 79
 QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120
 DB 80 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 139
 QY 121 PLVQECMWHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPFFAAIYRRH 180
 DB 140 PLVQECMWHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPFFAAIYRRH 199
 QY 181 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 240
 DB 200 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 259
 QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 300
 DB 260 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 319
 QY 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 360
 DB 320 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 379
 QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 411
 DB 380 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 430

RESULT 29

AAP91886
 ID AAP91886 standard; protein; 431 AA.

XX AC AAP91886;

DT 25-MAR-2003 (updated)
 DT 31-OCT-2002 (updated)
 DT 16-APR-1990 (first entry)

XX Sequence of prourokinase.

XX Low mol. wt. plasminogen activator; prourokinase; dissolve blood clot.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Protein 21..431
 FT /note= "Mature prourokinase."
 FT Region 170..179
 FT /note= "Preferred initiation region for the low mol. wt.
 FT plasminogen activators."

XX EP316068-A.

XX PD 17-MAY-1989.

XX

PF 07-OCT-1988; 88EP-0309417.

XX 09-OCT-1987; 87US-0107370.

PR 27-SEP-1988; 88US-0248727.

XX (COLB) COLLABORATIVE RES INC.

XX Mao JI;

XX WPI; 1989-146601/20.

XX N-PSDB; AAN91740.

XX Modified low mol. wt. plasminogen activator- formed of amino acids
 XX comprising the amino acid portion of prourokinase from 150 to 411

XX Fig 1; Page -; 27pp; English.

XX A low mol. wt. polypeptide plasminogen activator (PA) is claimed which is
 XX formed from amino acids 150-411 of prourokinase. The preferred initiation
 XX region for the low mol. wt. PA is indicated (see FT). A low mol wt. PA
 XX can be injected into blood in the body in vivo to dissolve clots
 XX without harm.
 XX (Updated on 31-OCT-2002 to add missing OS field.)
 XX (Updated on 25-MAR-2003 to correct PR field.)
 XX (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 431 AA;

Query Match 99.9%; Score 2298; DB 10; Length 431;
 Best Local Similarity 99.8%; Pred. No. 9.9e-178;
 Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 60

DB 21 SNELHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRG 80

QY 61 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 120

DB 81 KASTDTMGRPCLPWNSATVLOQTYHAHRS DALQLGLGKHNYCRNPDNRRPWCYVQVGLK 140

QY 121 PLVQECMWHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPFFAAIYRRH 180

DB 141 PLVQECMWHDCADGKKPSPPEELKFCQGGKTLRPRFKIIGGEFTTIENQPFFAAIYRRH 200

QY 181 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 240

DB 201 RGGSVTYVCGSLISPCWVISATHCFIDYPKKEDYIVLGRSLNSNTQGMKFEVENLI 260

QY 241 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 300

DB 261 LHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCETIGFGK 320

QY 301 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 360

DB 321 ENSTDYLYPEQLKMTVVKLI SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGGGL 380

QY 361 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 411

DB 381 VCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHTKENGIAL 431

RESULT 30

AAP94764

ID AAP94764 standard; protein; 431 AA.

XX AC AAP94764;

XX 25-MAR-2003 (updated)

DT 27-JUN-1990 (first entry)

XX Non-glycosylated prourokinase.

DE Prourokinase; CGE 195; plasminogen activator; blood clot lysis;

XX

Search completed: December 3, 2003, 14:39:12
Job time : 67.7422 secs

```
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Region 1..20
XX FT /label= signal sequence
XX
XX EP299706-A.
XX
XX PD 18-JAN-1989.
XX
XX PF 12-JUL-1988; 88EP-0306334.
XX
XX PR 13-JUL-1987; 87US-0072426.
XX PR 29-JUN-1988; 88US-0211279.
XX
XX PA (COLB ) COLLABORATIVE RES INC.
XX
XX PI Baltimore D, Moir DT, Broeze RJ;
XX
XX WPI; 1989-017204/03.
XX N-PSDE; AAN93079.
XX
XX New non-glycosylated, secreted plasminogen activator - pref. with
XX asparagine replaced or deleted, useful for treating blood clots,
XX expressed in non-mammalian cells.
XX
XX Disclosure; Page -: pp; English.
XX
XX myocardial infarction.
XX DNA encoding the protein was sequenced from plasmid pCGE195, a subclone
XX of two inserts isolated by screening a cDNA library prepd. from kidney
XX cell RNA. One of the original inserts, clone CGP31 (tag c) started in
XX the middle of the signal sequence. Mutants of the sequence, pref. in
XX which gcc (Ala) replaces aat (Asn) at nucleotides 1002-1004 (residue 302)
XX are used to transform hosts for the prodn. of non-glycosylated
XX prourokinase.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 431 AA;
XX
Query Match 99.9%; Score 2298; DB 10; Length 431;
Best Local Similarity 99.8%; Pred. No. 9.9e-178;
Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNEHQVPSNCDCNLGGTCVSNKIFSNHWCNCPKFKPGQGHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNEHQVPSNCDCNLGGTCVSNKIFSNHWCNCPKFKPGQGHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLFWSNATVLQQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLFWSNATVLQQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVOVGLK 140
QY 121 PLVQECWHDGADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 180
Db 141 PLVQECWHDGADGKPPSPPELKFQCGQKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
QY 181 RGGSVTVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFVEVNI 240
Db 201 RGGSVTVCGGSLISPCWISATHCFIDYPKKEDYIVLGRSLNSNTGEMKFVEVNI 260
QY 241 LHKDYSADTLAHHNDIALKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCEITGFGK 300
Db 261 LHKDYSADTLAHHNDIALKIRSKGRCQAPSRITQICLPSMYNDPQFGTSCEITGFGK 320
QY 301 ENSTDVLYPEQLKMTVVKLIISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDGGGGL 360
Db 321 ENSTDVLYPEQLKMTVVKLIISHRECOQPHYGSEVTTKMLCAADPQWKTDSCQDGGGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPWIRSHTKENGIAL 411
Db 381 LCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPWIRSHTKENGIAL 431
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 21.2586 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-3
Perfect score: 2301
Sequence: 1 SNELHQVPSNCDCLNGTGV.....VSHFLPWIRSHTKRENGIAL 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2298	99.9	431	1 UKHU	u-plasminogen acti
2	2148	93.4	433	1 UKBAY	u-plasminogen acti
3	1859.5	80.8	442	1 UKPG	u-plasminogen acti
4	1767	76.8	433	1 JN0560	u-plasminogen acti
5	1690.5	73.5	432	1 S18932	u-plasminogen acti
6	1660.5	72.2	433	1 UKMS	u-plasminogen acti
7	1004	43.6	434	1 A35005	u-plasminogen acti
8	864.5	37.6	477	2 JS0598	t-plasminogen acti
9	864.5	37.6	562	1 UKHUT	t-plasminogen acti
10	863.5	37.5	477	1 A34369	t-plasminogen acti
11	858.5	37.3	431	2 JS0599	t-plasminogen acti
12	844.5	36.7	559	1 A35029	t-plasminogen acti
13	836.5	36.4	477	2 JS0597	t-plasminogen acti
14	831.5	36.1	559	1 A29941	t-plasminogen acti
15	752	32.7	334	2 JS0600	t-plasminogen acti
16	735.5	32.0	655	1 A46888	hepatocyte growth
17	718.5	31.2	603	2 S28941	coagulation factor
18	692	30.1	615	1 KFHU12	coagulation factor
19	674.5	29.3	558	2 JC5878	plasma hyaluronan-
20	661.5	28.7	560	1 JC4795	plasma hyaluronan-
21	642	27.9	593	2 S45281	coagulation factor
22	507.5	22.1	460	2 B16245	plasmin (EC 3.4.21
23	504.5	21.9	790	1 PLPG	plasmin (EC 3.4.21
24	501	21.8	810	1 PLHU	plasmin (EC 3.4.21
25	497	21.6	810	2 B30848	plasmin (EC 3.4.21
26	497	21.6	812	1 PLMS	plasmin (EC 3.4.21
27	484.5	21.1	812	1 PLBO	plasmin (EC 3.4.21
28	484	21.0	1420	2 A32869	apolipoprotein(a)
29	474	20.6	4548	1 S00657	apoprotein(a) (EC

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminog
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form
C;Species: Homo sapiens (man)
C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastiao, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A;Title: The human urokinase-plasminogen activator gene and its promoter.
A;Reference number: A00931; MUID:85215647; PMID:2987867

ALIGNMENTS

30	457.5	19.9	455	2	A61545	plasmin (EC 3.4.21
31	445	19.3	761	2	JC5759	brain-specific ser
32	435	18.9	810	2	I46260	plasmin (EC 3.4.21
33	410	17.8	638	1	QMSPL	plasma kallikrein
34	408.5	17.8	343	1	A57014	proctasin (EC 3.4.
35	406	17.6	711	1	A47136	macrophage-stimula
36	402.5	17.5	417	1	S00845	hepsin (EC 3.4.21.
37	402	17.5	638	1	KQHUP	plasma kallikrein
38	401	17.4	416	1	J33777	hepsin (EC 3.4.21.
39	401	17.4	855	2	JC7731	membrane-bound arg
40	395.5	17.2	248	2	S55066	trypsin (EC 3.4.21
41	394.5	17.1	229	1	TRBOTR	trypsin (EC 3.4.21
42	394.5	17.1	716	1	JC5061	trypsin (EC 3.4.21
43	391.5	17.0	716	1	A40352	macrophage-stimula
44	390.5	17.0	247	2	S13813	trypsin (EC 3.4.21
45	390	16.9	263	2	A21195	chymotrypsin (EC 3

A;Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298

Query Match 93.4%; Score 2148; DB 1; Length 433;
Best Local Similarity 92.5%; Pred. No. 3.6e-158;
Matches 383; Conservative 17; Mismatches 10; Indels 4; Gaps 2;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGLNGGTCMSNKYFSSIHWCNCPKFGGQHCIDSKSKTCYEGNGHFYRG 79

QY 61 KASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVGLK 120
DB 80 KASTDTMGRSCLAWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVGLK 139

QY 121 PLVQECWHDCAKPKSPPEBELKFCQGGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 180
DB 140 QRYOECWHDCAKPKSPPEBELKFCQGGOKTLRPRFKIIGGFTTIENQFWFAAIYRRH 199

QY 181 RGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVLGRSLNNTQGMKFEVENLI 240
DB 200 RGSVTVYVCGSLISPCWVVSATHCFINYPKKEDYIVLGRSLNNTQGMKFEVENLI 259

QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAPSRITQICLPSMYNDPQ---FGTSCBITG 297
DB 260 LHEDYSADTLAHNDIALKIRSKGRCQAPSRITQICLPSMYNDPDPFGTSCBITG 319

QY 298 FGKENTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKTSCQDGS 357
DB 320 FGKENTDYLYPEQLKMTVVKLVSHQKCOQPHYGVSEVTTKMLCAADPQWKTSCQDGS 379

QY 358 GPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHKTEENGLAL 411
DB 380 GPLVCSIQGHMTLTGIVSGRGCAKDKPGVYTRVSRFLPWIRSHTRQENGLAL 433

RESULT 3
UKRG
u-plasminogen activator (EC 3.4.21.73) precursor - pig
N:Alternate names: uPA
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998
C:Accession: A00932
R:Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, B.
Nucleic Acids Res. 12, 9525-9541, 1984
A:Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A:Reference number: A00932; MUID:85087954; PMID:6096832
A:Accession: A00932
A:Molecule type: DNA
A:Residues: 1-240, 'H', 242-442 <NA>1
A:Experimental source: kidney cell line LLC-PK1
R:Nagamine, Y.
submitted to the Protein Sequence Database, December 1986
A:Reference number: A37566
A:Contents: annotation; correction to residue 241
C:Genetics:
A:Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
A:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-188/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F:72-153/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F:190-430/Domain: trypsin homology <TRY>
F:152/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F:235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 80.8%; Score 1859.5; DB 1; Length 442;
Best Local Similarity 79.3%; Pred. No. 6.7e-136;
Matches 334; Conservative 33; Mismatches 43; Indels 11; Gaps 2;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHFY 58
DB 21 SREL-QVPSDCGLNGGTCMSNKYFSSIHWCNCPKFGGQHCIDSKSKTCYEGNGHFY 79

DB 21 SHELHQPNSCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHYS 80

QY 59 RGAASDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG 118
DB 81 RGAANTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG 140

QY 119 LKPLVQECWHDCA-----DGKKPSPPEBELKFCQGGOKTLRPRFKIIGGFTTIEN 169
DB 141 LKQLVQECWVPCNSGGESHRRPAYDGKPFSTPEKVFQCGQKALRPRFKIVGKSTTIEN 200

QY 170 QPWFAAIYRRHSGSVTVYVCGSLISPCWVVSATHCFIDYPKKEDYIVLGRSLNNTQ 229
DB 201 QPWFAAIYRRHSGSVTVYVCGSLISPCWVVSATHCFINYPKKEDYIVLGRSLNNTQ 260

QY 230 GEMKFEVENLIHKDYSADTLAHNDIALKIRSKGRCQAPSRITQICLPSMYNDPQF 289
DB 261 GEMKFEVENLIHKDYSADTLAHNDIALKIRSKGRCQAPSRITQICLPSMYNDPQF 320

QY 290 GTSCEITGFGKENTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKT 349
DB 321 GASCEIVGFKEDPSDYLYPEQLKMTVVKLVSHRECOQPHYGVSEVTTKMLCAADPQWKT 380

QY 350 DSCQDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSHFLPWIRSHKTEENGL 409
DB 381 DSCQDSGGPLVCSLQGRMTLTGIVSGRGCAKDKPGVYTRVSRFLPWIRSHKTEENGL 440

QY 410 A 410
DB 441 A 441

RESULT 4
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N:Alternate names: uPA
C:Species: Bos primigenius taurus (cattle)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C:Accession: JN0560
R:Kraatzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A:Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and indi
A:Reference number: JN0560; MUID:93216119; PMID:8385052
A:Accession: JN0560
A:Molecule type: mRNA
A:Residues: 1-433 <KRA>
A:Cross-references: GB:I03546; NID:gl63800; PIDN:AAA51419.1; PID:gl63801
C:Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C:Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-179/Product: plasminogen activator chain A #status predicted <MA1>
F:21-179/Product: urokinase-type plasminogen activator chain A #status predicted <MA1>
F:33-64/Domain: EGF homology <EGF>
F:72-153/Domain: kringle homology <KRG>
F:181-433/Product: plasminogen activator chain B #status predicted <MA2>
F:181-421/Domain: trypsin homology <TRY>
F:170-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F:226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 76.8%; Score 1767; DB 1; Length 433;
Best Local Similarity 75.1%; Pred. No. 9e-129;
Matches 310; Conservative 45; Mismatches 56; Indels 2; Gaps 1;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHFY 58
DB 21 SNEVHKGESGCGCLNGGTCVYFYFNIHWCNCPKFGGQHCIDSKSKTCYEGNGHYS 80

QY 59 RGAASDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG 118
DB 81 RGAANTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNCRNPNRRPWCYVOVG 140

QY 119 LKPLVQECWHDCAKPKSPPEBELKFCQGGOKTLRPRFKIIGGFTTIENQFWFAAIYR 178
DB 141 LKQFVQFCWQDCSVGKSPSPPEBELKFCQGGOKTLRPRFKIIGGFTTIENQFWFAAIYR 200

Db 150 HDCSLKPPSSVDQGFQCGQKALRPRFKIVGGFEVEVQWFAAIYQKNGKGGSPSP 209
QY 188 VCGSLSPCWVWISATHCFIDYPKEDYIVYLGSRRLNSNTQGMKPFVENLILHKOYSA 247
Db 210 KCGSLSPCWVWISATHCFIDYPKEDYIVYLGSRRLNSNTQGMKPFVENLILHKOYSA 247
QY 248 DTLAHHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKNSDYL 307
Db 270 DSLAYHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKNSDYL 329
QY 308 YPEOLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGGLVCSLQGR 367
Db 330 YPKNLKMSVVKLVSHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGGLVCSLQGR 367
QY 368 MTLGIVSWGRCALKDKPGVTVVSHFLPWIRSHTEENGLA 410
Db 390 PTLGIVSWGRCALKDKPGVTVVSHFLPWIRSHTEENGLA 432

RESULT 7
A35005
u-plasminogen activator (EC 3.4.21.73) precursor - chicken
N;Alternate names: uPA
C;Species: Gallus gallus (chicken)
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 16-Jul-1999
C;Accession: A35005
R;Leslie, N.D.; Kessler, C.A.; Bell, S.M.; Degen, J.L.
J. Biol. Chem. 265, 1339-1344, 1990
A;Title: The chicken urokinase-type plasminogen activator gene.
A;Reference number: A35005; MUID:90110185; PMID:2295632
A;Accession: A35005
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-434 <LRS>
A;Cross-references: GB:J05187; NID:g212858; PIDN:AAA49131.1; PID:g212859
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-171/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F;40-71/Domain: EGF homology <EGF>
F;79-158/Domain: kringle homology <KRG>
F;173-428/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F;173-416/Domain: trypsin homology <TRY>
F;162-296,202-218,210-285,310-379,342-358,369-397/Disulfide bonds: #status predicted
F;217,272,373/Active site: His, Asp, Ser #status predicted

Query Match 43.6%; Score 1004; DB 1; Length 434;
Best Local Similarity 47.4%; Pred. No. 6.6e-70;
Matches 188; Conservative 64; Mismatches 123; Indels 22; Gaps 6;

QY 11 CDCLANGTCVSNKYFSNIHWCNCPKFGQHCIEDSKTCYEGNGHFYRGKASTDTWGRP 70
Db 40 CQCLNGTCTIYRFFSQIKCLCEGGLGHLCEIDTNSICYSGNGEDYRGNAEDP----G 95
QY 71 CLPWNSATVLQ-QTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVH 129
Db 96 CLYWDHPSVIRWGDYHADLNALQGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVH 149
QY 130 DCADGKPPSPPELKCQCKTLRPRFKIIGFEFTTIENQPPAAIYRHRGSGVTVVC 189
Db 150 -----TPCSTIEKCECTQGRSFYKFKIVGGSQAQVETQPMIAGIFQNM-GTDQFLC 202
QY 190 GGSILSPCWVWISATHCFID----YPKEDYIVYLGSRRLNSNTQGMKPFVENLILHKOY 245
Db 203 GGSILDPQWVLTAAHCYNTKQPNKSVTKVFLGKSLINTNDEHQVFWDEIISHPDF 262
QY 246 SADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEITGFGKNSD 305
Db 263 TDHTGGNDNDIALIRINTASQCAVESYVTVCLPEKNLNDYNTWCEIAGYKQNSYD 322
QY 306 YLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWTKDSQCGSGGLVCSLQ 365

Db 323 IYYAQLMSATVNLISQDDCKNKYYDSTRVTDNMVACGDLWETDACKGDSGPMVCEHN 382
QY 366 GRMTLTGIVSWGRCALKDKPGVTVVSHFLPWIRSH 402
Db 383 GRMTLYGIVSWGRCALKDKPGVTVVSHFLPWIRSH 419

RESULT 8
JS0598
t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0598
R;Kraetschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat D
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0598
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359
F;185,398/Binding site: carboxylate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 37.6%; Score 864.5; DB 2; Length 477;
Best Local Similarity 42.9%; Pred. No. 4.2e-59;
Matches 178; Conservative 60; Mismatches 148; Indels 29; Gaps 9;

QY 3 ELHQVP----SNCDCLNGTCTVSNKYFSNIHWCNCPKFGQHCIEDSKTCYEGNGHFY 58
Db 78 QCHTVPVKSCSELRCFNGTCTVSNKYFSNIHWCNCPKFGQHCIEDSKTCYEGNGHFY 136
QY 59 RGKASTDTMGRPCLPWNSATVLQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGL 118
Db 137 RGTWSTSEGAQCINMNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNRRPWCYVQVGL 196
QY 119 LKPLVQECMVHDCADGKPPPEELKFCQG-QKTLRPRFKIIGFEFTTIENQPPAAIY 177
Db 197 SKFILEFCSPVPCS-----KATGLRKYKPEQLHSTGGLFTDITSHPWQAIF 244
QY 178 RHRGGS-VTVVCGGSLISPCWVWISATHCFID-YPKEDYIVYLGSRRLNSNTQGMKFE 235
Db 245 AQNRSSGERFLCGGILISSCWLTAAHCYRPPHQLRVV-LGTRYVYKPGKEQTTFE 303
QY 236 VENLILHKOYADTLAHHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPOFGTSCEI 295
Db 304 VEKCIHVEEEDDT--YNNDIALLQLKSGSPQCAQESDVSRAICLPEANLQLPDWTECEL 361
QY 296 TGFGENKSTLYPEOLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWTK----- 349
Db 362 SGYKHKSSPFFSBEQLKEGHVLPSSRCTSKFLNKTNNMLCAGDTRSGEIPNVH 421
QY 350 DSCQSGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVTVVSHFLPWIRSHTK 404
Db 422 DACQSGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVTVVSHFLPWIRSHTK 476

RESULT 9
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N;Alternate names: t-PA; tissue plasminogen activator

C;Species: Homo sapiens (man)
C;Date: 14-Nov-1993 #sequence_revision 14-Nov-1993 #text_change 08-Dec-2000
C;Accession: A94004; A23529; J0562; A93293; S02125; A91343; A93951; A91322; A54645; I60
R;NY, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation
A;Reference number: A94004; MUID:84298137; PMID:6089198
A;Accession: A94004
A;Molecule type: DNA
A;Residues: 1-562 <NNT>
A;Cross-references: GB:L00141
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
R;Friezeher Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A;Title: The human tissue plasminogen activator gene.
A;Reference number: A23529; MUID:86196143; PMID:3009482
A;Accession: A23529
A;Molecule type: DNA
A;Residues: 1-562 <DEG>
A;Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuoka, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A;Title: Purification and characterization of tissue plasminogen activator secreted by h
A;Reference number: J0562; MUID:91291340; PMID:1368681
A;Accession: J0562
A;Molecule type: mRNA
A;Residues: 31-562 <ITA>
A;Cross-references: DBJ:D01096; NID:G220128; PIDN:BA000881.1; PID:G441174
A;Experimental source: embryonic lung fibroblast IMR-90 cells
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A;Reference number: A93293; MUID:83115262; PMID:6337343
A;Accession: A93293
A;Molecule type: mRNA
A;Residues: 1-562 <PEN>
A;Cross-references: GB:L00141
A;Experimental source: melanoma cells
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A;Reference number: S02125; MUID:88262579; PMID:3133640
A;Accession: S02125
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-562 <SAS>
A;Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244
A;Experimental source: fetal lung cells
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
FEBS Lett. 189, 145-149, 1985
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A;Reference number: A91343; MUID:85285620; PMID:3896853
A;Accession: A91343
A;Molecule type: mRNA
A;Residues: 1-38, 'G', 86-433, 'E', 435-562 <XAG>
A;Experimental source: Detroit 562 cells; ATCC 138
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
A;Reference number: A93951; MUID:83169656; PMID:6572897
A;Accession: A93951
A;Molecule type: mRNA
A;Residues: 251-358 <EDL>
A;Experimental source: melanoma cells
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.
A;Reference number: A90488; MUID:85000468; PMID:6433976
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active and
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984

A;Title: Differences between uterine and melanoma forms of tissue plasminogen activato
A;Reference number: A91322; MUID:84158956; PMID:6538514
A;Accession: A91322
A;Molecule type: protein
A;Residues: 33-45;311-320 <POH>
A;Experimental source: uterus
A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A;Reference number: A37567; MUID:87033611; PMID:3021732
A;Contents: annotation; fibrin binding site
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Eng
EMBO J. 5, 3525-3530, 1986
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen
A;Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type
A;Reference number: A60902; MUID:89044681; PMID:3142086
A;Contents: annotation; novel forms of expressed recombinant t-PA
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.
Mol. Biol. Med. 3, 279-292, 1986
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its ex
A;Reference number: A54645; MUID:86284200; PMID:3090401
A;Accession: A54645
A;Molecule type: mRNA
A;Residues: 1-562 <HAR>
A;Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
A;Note: parts of this sequence were confirmed by peptide sequencing
R;Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A;Reference number: I60110; MUID:88054470; PMID:2824147
A;Accession: I60110
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-562 <RES>
A;Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A;Title: Isolation and characterization of the human tissue-type plasminogen activator
A;Reference number: I55232; MUID:85289338; PMID:3161893
A;Accession: I55232
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-36 <RE2>
A;Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a singl
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C;Genetics: GDB:PLAT
A;Gene: GDB:119496; OMIM:173370
A;Cross-references: GDB:119496; OMIM:173370
A;Map position: 8p12-8p12
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 5
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Superfamilies: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F1-23/Domain: signal sequence #status predicted <SIG>
F1-23/Domain: signal sequence #status predicted <PRO>
F1-23/Domain: propeptide #status predicted <PRO>
F33-310/Product: t-plasminogen activator #status experimental <NAT>
F33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F41-78/Domain: fibronectin type I repeat homology <IF1>
F41-78/Domain: EGF homology <EGF>
F46-119/Domain: kringle homology <KR1>
F127-208/Domain: kringle homology <KR2>
F215-296/Domain: kringle homology <KR2>
F311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F311-566/Domain: trypsin homology <TRY>
F41-71-69-78-86-97-91-108-110-119-127-208-148-190-179-203-215-236-236-278-291-299
F152-483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F219/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F357, 406/Active site: His, Asp #status predicted

F:513/Active site: Ser #status experimental

Query Match 37.6%; Score 864.5; DB 1; Length 562;
Best Local Similarity 38.1%; Pred. No. 5.1e-59;
Matches 196; Conservative 56; Mismatches 157; Indels 89; Gaps 11;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNSLHWCNCPKFGQHCEDKSKTCYEGNGHFY 58
DB 77 QCHSVFVKSCSEPCFCNGGTCQOALYFSDP-VQCQPEGFAGKCCEDTRATCYEDQGISY 135
QY 59 RGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRRWCYQVVG 118
DB 136 RGTWSTAESGAECNTNWSALAQKPSGRPPDAIRLGLGNHNYCRNPDRDRSKPWCYVFA 195
QY 119 LKPLVOECMVHDCADG-----KQPS 138
DB 196 GKYSSFECSFPAACSEGNDCYFGNGSAYRGTHSLITSGASCLPWNMSMILIGKVYTAQNP 255
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKLIGGE 163
DB 256 AQALGLGKHNYCRNPDPGDAKPWCHLVKNRRLTWECYCDVPCSTCGLRQYQSQPFRKGG 315
QY 164 FTTIENQPFALYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDVIVILGRS 222
DB 316 PADIASHPQWAAIFAKHRSFGERFLCGGILISSCWILSAHCFQERFPFPHLTVILGRT 375
QY 223 RLNSNTQGENKFEVENILHKOYSADTLAHNDIALLKIRSKRGCRCQAQPSRTIQTICLPS 282
DB 376 YRVVPDEEQKFEKVIYHKEFDDDT--YDNDIALQLKSDSRCAQESSVVRVTCCLPP 433
QY 283 MYNDPQPGTSCETLGFGENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 342
DB 434 ADLQLPWTCELSGYGKHREALSPFYSERLKAHVRILYPSRCTSQHLLNRTVTDNMLCA 493
QY 343 AD-----POWKT-DSQCGSGGLPVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFL 396
DB 494 GDRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKQDVGQVTKVNYL 553
QY 397 PWIRSHTK 404
DB 554 DWIRDNR 561

RESULT 10
A34369
t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
C;Species: Megaderma lyra
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
R;Accession: A34369
R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jacob
J. Biol. Chem. 264, 17947-17952, 1989
A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasmin
A;Reference number: A34369; MUID:90036867; PMID:2509450
A;Accession: A34369
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-477 <GAR>
A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-477/Product: plasminogen activator #status predicted <PLA>
F:42-79/Domain: fibronectin type I repeat homology <IFA>
F:87-120/Domain: EGF homology <EGF>
F:128-209/Domain: kringle homology <KR>
F:226-471/Domain: trypsin homology <TRY>
F:42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
F:272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 37.5%; Score 863.5; DB 1; Length 477;
Best Local Similarity 42.9%; Pred. No. 5.1e-59;
Matches 178; Conservative 60; Mismatches 148; Indels 29; Gaps 9;

QY 3 ELHQVP-----SNCDCLNGGTCVSNKYFNSLHWCNCPKFGQHCEDKSKTCYEGNGHFY 58
DB 78 QCHTVFVKSCSELRCFCNGGTCQOAAFSDF-VQCQPKGYTGKQCEVDTHATCYADQGVTY 136
QY 59 RGKASTDTWGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRRWCYQVVG 118
DB 137 RGTWSTSEGAQCINNWSNLLTERTYNGRRSDAILGLGNHNYCRNPDRNNSKPCYVKA 196
QY 119 LKPLVOECMVHDCADGKPPSPPEELKFCQG-QKTLRPRFKIIGGEFTTIENQPFPAIY 177
DB 197 SKFILEFCSPVCS-----KATGLRKYKEPQLHSTGGLFTDITSHPWQAIF 244
QY 178 RHRGGS-VTVYVCGGSLISPCWVISATHCFID-YPKKEDVIVYVGLSRSLNSNTQGENKFE 235
DB 245 AQNRSSGERFLCGGILISSCWVLTAAHCFQERYPPQHLRV-LGRYRVKPGKEQTPE 303
QY 236 VENLLIHKYSADTLAHNDIALLKIRSKRGCRCQAQPSRTIQTICLPSMYNDPQFGTSC 295
DB 304 VEKCIVHEPDDDT--YNDIALQLKSGSPQCAQESDVRAICLPEANLQLPDWTECEL 361
QY 296 TGRGENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPQWKT----- 349
DB 362 SGYGHKSSPFPYSEQLKEGHVRLYPSRCTSKFLNKTVTKMLCAGDTRSGEIHENVH 421
QY 350 DSCQSGGGLPVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHPLPWRSHTK 404
DB 422 DACQSGGGLPVCRNDNHHMTLLGLISWVGCGEKDIPGVTVKTVNYLGMWRDNR 476

RESULT 11

JS0599
t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
R;Accession: JS0599
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat D
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0599
A;Molecule type: mRNA
A;Residues: 1-431 <KRA>
A;Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-431/Product: plasminogen activator beta #status predicted <PLA>
F:41-74/Domain: EGF homology <EGF>
F:82-163/Domain: kringle homology <KR>
F:180-425/Domain: trypsin homology <TRY>
F:41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/[Disulfide b
F:139,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:179-180/Cleavage site: His-Ser (plasma) #status predicted
F:226,275,382/Active site: His, Asp, Ser #status predicted
F:345-361,378-406/Disulfide bonds: #status predicted

Query Match 37.3%; Score 858.5; DB 2; Length 431;
Best Local Similarity 43.2%; Pred. No. 1.1e-58;
Matches 175; Conservative 59; Mismatches 146; Indels 25; Gaps 8;

QY 9 SNCDCLNGGTCVSNKYFNSLHWCNCPKFGQHCEDKSKTCYEGNGHFYRGKASTDTMG 68
DB 42 SELRCFCNGGTCQOAAFSDF-VQCQPKGYTGKQCEVDTHATCYADQGVTVRGVSTSESG 100
QY 69 RPLCPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRRWCYQVGLKPLVQECW 128
DB 101 AQCNWNSNLLTERTYNGRRSDAILGLGNHNYCRNPDRNNSKPCYVKAISKFLFECV 160
QY 129 HDCADGKKPSPPEELKFCQG-QKTLRPRFKIIGGEFTTIENQPFPAIYRRHRGGS-VT 186

Db 161 PVCS-----KATCGLRKYPQLHSTGGFTDITSHPWQAALFAQNRSSGER 208
QY 187 YVCGSLISPCWVISATHCFID-YPKKEDYIVVLGRSLNSNTQGMKFEVENLILHKDY 245
Db 209 FLCCGGLISCCWLTAHCFQERYPPQHLRVV-LGRTYRVKPKQETFEVCKIIEEF 267
QY 246 SADTLAHNDIALLKIRSEKGRCAQSRRTIQTCLFSMYNDPQFGSCBITGKGNSTD 305
Db 268 DDDT--YNNIDIALQLKSGSPQAQESDSVRAICLPEANLQLPDWTECELSGVGKHKSS 325
QY 306 YLYPEQLKMTVVKLISHRECCQPHYYGSEVTTKMLCAADPQWKT-----DSCQGSGGP 359
Db 326 PFYSEQLKEGHVRLYSSRCTSKFLNKVTNNMLCAGDSRGEIYPNVHDCQGSGGP 385
QY 360 LVCSLQGRMLTIGVSMGRGKALKDKPGVYTRYVSHFLPWIRSHTK 404
Db 386 LVCWNDHMTLLGIISWVGCGEKDIPGVYTKVNYLIGWIRDNR 430
RESULT 12
A35029
t-plasminogen activator (EC 3.4.21.68) precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35029; A31597
J;Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 265, 2022-2027, 1990
A;Title: The structure of the TATA-less rat tissue-type plasminogen activator gene. Spec
A;Reference number: A35029; MUID:90130448; PMID:2105315
A;Accession: A35029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M31197; NID:G207429; PIDN:AAA42261.1; PID:G207431; GB:J05226
R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A;Reference number: A31597; MUID:89170114; PMID:3114845
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379, 'K', 381-559 <NYT>
A;Cross-references: GB:M23697; NID:G530159; PIDN:AAA41812.1; PID:G530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1FI>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68, 66-75, 83-94, 88-105, 107-116, 124-205, 145-187, 176-200, 213-294, 234-276, 265-289, 297-4
F;149, 481/Binding site: carboxylate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355, 404, 510/Active site: His, Asp, Ser #status predicted
Query Match 36.7%; Score 844.5; DB 1; Length 559;
Best Local Similarity 36.9%; Pred. No. 1.8e-57;
Matches 182; Conservative 64; Mismatches 150; Indels 97; Gaps 12;
QY 3 ELHQP-----SNCDLNGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYF 58
Db 74 QCHSVPRVSRSEPRCFNGGTCCQALYFSDF-VCQCPDGFVGKRCIDIDTRATCFEGQGIT 132
QY 59 RGKASTDTMGPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
Db 133 RGTWSTAENGACINWNSSALSQRPYSARRPNAIKLGLGNHYCRNPNDRDPKPCYFKA 192
QY 119 LKPLVQECM-----VHDCADGK-----KPSPP----- 141

Db 193 GKYTEFCSTPACPKGPTEDCYVGKGVTYRGTHSFTTSKASCLPWNMILGKTYTAWRA 252
QY 142 -----BELKFQ-----COQKTLR-PRFKIIG 162
Db 253 NSQALGLGRHNYCRNPDGDAKPMCHVMKDKLITWEYCDMSPCSTCGLRQYKQPFRIKGG 312
QY 163 EFTTIENQWFAAIY-RRIRGGSVTVCGLSLPCWVISATHCFIDYPKKEDYIVVLGR 221
Db 313 LFTDITSHPWQRAIFVKNKRSPGERFLCGVLSSCWLSSAAHCFVERFPFHHLKVVVLGR 372
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLKIRSEKGRCAQSRRTIQTICLP 281
Db 373 TYRVVGEQEFTFEIKYIVHKEFDDT--YNDIALQLRSSQCAQESSVGTACLP 430
QY 282 SMYNDPQF---CTSCBITGFGKENSTDYLPBQLKMTVVKLISHRECCQPHYYGSEVTT 337
Db 431 ---DPDVLQPDWTECELSYGKHEASPPFSDRLKEAHVRLYPSRCTSHLFNKTITS 486
QY 338 KMLCAADP-----QWKTDSCQDGGPLVCSLQGRMLTIGVSMGRGKALKDKPGVYTRY 392
Db 487 NMLCAGDTRTGGNQDVHDACQDGGPLVCMIDKRMTLGLIISWGLCGCGKDPFGIYTKV 546
QY 393 SHFLPWIRSHTK 405
Db 547 TNYLNIQDNMKQ 559
RESULT 13
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0597
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidel, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat D
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0597
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63987; NID:G166070; PIDN:AAA31591.1; PID:G166071
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KRG>
F;226-471/Domain: trypsin homology <TRY>
F;42-72, 70-79, 87-98, 92-109, 111-120, 128-209, 149-191, 180-204, 214-345, 257-273, 265-334, 359
F;153, 398/Binding site: carboxylate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272, 321, 428/Active site: His, Asp, Ser #status predicted
Query Match 36.4%; Score 836.5; DB 2; Length 477;
Best Local Similarity 42.2%; Pred. No. 6.1e-57;
Matches 175; Conservative 56; Mismatches 147; Indels 37; Gaps 10;
QY 5 HQVPSN-CD---CLNGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHYF 60
Db 80 HTVFNVCSEPRCFNGGTCCQALYFSDF-VCQCPAGITGKRCVDTTRATCYEGGVYTRG 138
QY 61 KASTDTMGPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 139 TWSTAERSEVCINWNSLLTRTYNGRMPDAFNGLGNHYCRNPNGAPKPCYVIKAG 198
QY 121 PLVQECMVHDCADGKPKSPPELKFQCG-QKTLRPRFKIIGGFTTIENOPFAAIYR 179
Db 199 FTSESCSVFVCS-----KATCGLRKYPQLHSTGGFTDITSHPWQAALFAQ 246
QY 180 HRGGS-VTVYVCGSLISPCWVISATHCFIDYPKKEDYI-----VYLGRSLNSNTQGMK 233

```

Db 247 NRSSGERFLCGGILLISSCWMLTAACFP-----QESYLPDLQKVLGRITVRVPGEEQT 301
QY 234 FEVENILHKDYGADTLAHHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPPQFTSC 293
Db 302 FKVKYKIVHKEFDDT--YNNDIALLQLKSDSPQCAQESVRAICLPEANLQLPDWTEC 359
QY 294 EITGFKENSTLYLPEQLKMTVVKLISHRECQOPHYGSEVTTKMLCAADPOWKT----- 349
Db 360 ELSGYGKHSSSPFYSEQLKEGHVRLVPSRCAPKFLFNKTVTNMMLCAGDTSGEITPN 419
QY 350 --DSCGDSGGLVCSLQGRMTLTGIVSGRCALKDKPGVTVRSHFLPWIRSH 402
Db 420 VHDACQDSGGPLVGVNDNHTLLGIISWVGGEKDPVGVYTKVNYLGIWRDN 474

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A29941; S48205; S48207; S48206
R:Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A:Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA
A:Reference number: A29941; MUID:88087303; PMID:2826484
A:Accession: A29941
A:Molecule type: mRNA
A:Residues: 1-559 <RIC>
A:Cross-references: GB:J03520; NID:g202109; PIDN:AAA40470.1; PID:g202110
R:Liijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A:Title: Characterization of the murine plasma fibrinolytic system.
A:Reference number: S48202; MUID:95010076; PMID:7523120
A:Accession: S48205
A:Molecule type: protein
A:Residues: 33-37, 'X', 39-40 <LIJ>
A:Accession: S48207
A:Molecule type: protein
A:Residues: 309-316 <LIJ>
A:Accession: S48206
A:Molecule type: protein
A:Residues: 33-37, 'X', 39-40 <LIW>
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-29/Domain: propeptide #status predicted <PRO>
F:30-559/Product: t-plasminogen activator #status predicted <MAT>
F:30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F:38-75/Domain: fibronectin type I repeat homology <1f1>
F:83-116/Domain: EGF homology <EGF>
F:124-205/Domain: kringle homology <KR1>
F:213-294/Domain: kringle homology <KR2>
F:309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F:309-553/Domain: trypsin homology <TRY>
F:368-66,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F:149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:308-309/Cleavage site: Arg-ile (plasmin, trypsin) #status predicted
F:355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 36.1%; Score 831.5; DB 1; Length 559;
Best Local Similarity 35.9%; Pred. No. 1.8e-56;
Matches 177; Conservative 65; Mismatches 154; Indels 97; Gaps 10;

QY 3 ELHQVP-----SNDCINGTCVSNKYFSNIHWCNCPKKGQGHCEIDKSKTCVEGNHGFY 58
Db 74 QCHSVPRVSRSEPRCFNGGTCCQALYFSDF--VCQCPDGFVGKRCIDITRATCFEEQITY 132
QY 59 RGRASDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRNPDNRRRRCWYQVG 118
Db 133 RGTWSTAESGAECINMNSVLSUKPYNARRPNAIKLGLGNHNYCRNPDRLKPKWYFKA 192
QY 119 LKPLVQECMVHDCADGKKPS----- 138

```

```

Db 193 GKYTECSTPACPKGSEDCYVGKGYRGTGTHSLTTSQASCLPWNISVLMGKSYTAWRT 252
QY 139 -----SPPEELK-----FQCG-QKTLRPRFKTIQG 162
Db 253 NSQALGLARHNYCRNPDGDARPWCHVKDKRKLWYECDSMSPCSTCGLRQYKRQFRIKGG 312
QY 163 EFTTIENQWFAAIY-RRHRGGSVTVVCGSLISPCWVISATHCFIDYPKKEIDYVVLGR 221
Db 313 LYTDITSHPMQAAIFVKNKESPGERFLCGGLVLISSCWVLSAAHCFLEFRFPNHLKVVLGR 372
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIQTICLP 281
Db 373 TYRVVGESEOTEIEKIYVHEEFDDT--YNDIALLQLRSQSKQAQESSVGTACLP 430
QY 282 SMYNDPQF-----GTSCEITGFGKENSTLYLPEQLKMTVVKLISHRECQOPHYGSEVTT 337
Db 431 ----DNLQLPDWTECELSGYKHEASSPFFSRLKEAHVRLYPSRRCTSHLFNKTVTN 486
QY 338 KMLCAADP-----QWKTDSCGDSGGPLVCSLOGRMTLTGIVSGRCALKDKPGVTVR 392
Db 487 NMLCAGDTRSGNQDLHDACQDSGGPLVCMINKQMTLTGIIISWGLCGQKDPVGVYTKV 546
QY 393 SHELPLWIRSHTE 405
Db 547 TNYLDWIHNMKQ 559

RESULT 15
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N:Alternate names: tissue plasminogen activator
C:Species: Desmodus rotundus (common vampire bat)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: JS0600
R:Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A:Title: The plasminogen activator family from the salivary gland of the vampire bat D
A:Reference number: JS0597; MUID:92039036; PMID:1937019
A:Accession: JS0600
A:Molecule type: mRNA
A:Residues: 1-394 <KRA>
A:Cross-references: GB:M63990; NID:gl66078; PIDN:AAA31595.1; PID:gl66079
A:Note: the authors translated the codon ATC for residue 75 as Thr
C:Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-36/Domain: propeptide #status predicted <PRO>
F:37-394/Product: plasminogen activator gamma #status predicted <PLA>
F:45-126/Domain: kringle homology <KR>
F:143-388/Domain: trypsin homology <TRY>
F:45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds
F:142-143/Cleavage site: His-Ser (plasmin) #status predicted
F:189,238,345/Active site: His, Asp, Ser #status predicted
F:315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.7%; Score 752; DB 2; Length 394;
Best Local Similarity 42.0%; Pred. No. 1.6e-50;
Matches 155; Conservative 56; Mismatches 134; Indels 24; Gaps 7;

QY 45 DKSKTCEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSALQLGLGKHNYCRN 104
Db 40 DPATCYKQGVTVRGTWSTESGAQCINWNSNLLIRTYNGRMPKAVRGLGNHNYCRN 99
QY 105 PDNRRRRCWYQVGLKPLVQECMVHDCADGKKPSPPPEELKFCQG-QKTLRPRFKTIIGE 163
Db 100 PDGASKEPCVIVKARKFTSSCSVPVCS-----KATCGLRKYKEPQLHSTGGL 147
QY 164 FTTIENQWFAAIYRHRGGS-VTVVCGSLISPCWVISATHCFID-YPKKEIDYVVLGR 221
Db 148 FTDITSHPMQAAIFVKNRNRSSGERFLCGGLIISCWVLTAAHCFQRYPPQHLRVV-LGR 206
QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHHNDIALKIRSKGRCQAQPSRTIQTICLP 281

```

```
Db 207 TYRKPGEQTEVERKCIVHEEFDDDT--YNNDIALLQLKSGSPQCAQESDSVRAICLP 264
Qy 282 SMYNDPQFGTSCETITGKKNSTDYLPQELKMTVVKLISHRECCQPHYGVSEVTTKMLC 341
Db 265 EANIQLPDWTECELSGKHKSSFFSEQLKEGHVLYSSRSTKSFELNKTVTNNMLC 324
Qy 342 AADPQWKT-----DSQSGSGPLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHF 395
Db 325 AGDTRSGEIVPNVHDAQCGSGGGLVCMNDNHNMTLLGIISGVGCGEKDIPGVYTKVTNY 384
Qy 396 LPWIRSHTK 404
Db 385 LGWIRDNR 393

RESULT 16
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R: Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine protease
d coagulation factor XII.
A:Reference number: A46688; MUID:93252878; PMID:7683665
A:Accession: A46688
A:Molecule type: mRNA
A:Residues: 1-655 <M>Y>
A:Cross-references: DDJ: D14012; NID: g219680; PIDN: BAA0113.1; PID: g219681
A:Experimental source: liver (mRNA); serum (protein)
A:Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBI:131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains, c
A:Genetics:
A:Gene: GDB: HGPAC; HGPA; HGPAP
A:Cross-references: GDB: 9954514
A:Map position: 4p16-4p16
C:Function:
A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:108-148/Domain: fibronectin type II repeat homology <1F2>
F:164-197/Domain: EGF homology <EG1>
F:202-237/Domain: fibronectin type I repeat homology <1F1>
F:245-278/Domain: EGF homology <EG2>
F:286-367/Domain: kringle homology <KRG>
F:373-407/Product: hepatocyte growth factor activator light chain #status experimental <
F:408-655/Product: hepatocyte growth factor activator heavy chain #status experimental <
F:408-641/Domain: trypsin homology <TRY>
F:40,48,290,469,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36
F:447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.0%; Score 735.5; DB 1; Length 655;
Best Local Similarity 37.6%; Pred. No. 5.4e-49;
Matches 160; Conservative 53; Mismatches 166; Indels 47; Gaps 9;

Qy 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPKFGGQHCEIDKSKTCYEGNGH 56
Db 242 HTACLSFLNGGTC-----HLIVATGTVTACPPGAGRLNLEPDERCFLNGT 292
Qy 57 FYRGKASTDTWGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDNRPRWCYVQ 116
Db 293 GYRGVASTASGLSCLAWSNLLYQELHVDVSGAALLGLGPHAYCRNPDNDRPWCYV 352
Qy 117 VGLKPLVCEMVHDC-----ADGKPSPEELKFCQGK-----TLRPRFX 158
Db 353 KDSALSWEYCLEACESLTRVOLSPDLLATLPEPASGRQ---ACGRHKKRTFLRPR-- 407
Qy 159 IIGGEFTTIENQPFPAAYRRHRGSGSVTVCGGSLISPCWVISATHCFIDYPKKEDYIV 218
```

```
Db 408 IIGSSSLPGSHPLAAIY---IGDS---FCAGSLVHTCMWVSAHCFSPPRDSVSV 461
Qy 219 LGRSLRNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCACQPSRTTQTI 278
Db 462 LGQHFRRTDVTDTGFGIEKYIPYTLYSVFNPSDH-DLILRLKKKGDRCATRFQVPI 520
Qy 279 CLPSMYNDPQFGTSCETITGKKNSTDYLPQELKMTVVKLISHRECCQPHYGVSEVTTK 338
Db 521 CLPBPGGSTFPAGHKCQIAGWGHLDENVSGYSSSLREALVPLVADHKCSSPPEYVADISPN 580
Qy 339 MLCRAADPQWKTDCQSGSGGGLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHF 398
Db 581 MLCAGYFDCKSDACQSGSGGLPLACRNGVAYLGIISWGDGCGRLHKPGVYTRVANYVDV 640
Qy 399 IRSHTK 404
Db 641 INDIR 646

RESULT 17
S28941
coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)
N: Alternate names: Hageman factor
C: Species: Cavia porcellus (guinea pig)
C: Date: 25-Feb-1994 #sequence_revision 03-Aug-1995 #text_change 21-Jan-2000
C: Accession: S28941
R: Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe,
Biochim. Biophys. Acta 1159, 113-121, 1992
A: Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage
A: Reference number: S28941; MUID: 93003367; PMID: 1390917
A: Accession: S28941
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-603 <SM>
A: Cross-references: EMBL: X68615; NID: g49578; PIDN: CAA48600.1; PID: g49579
A: Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C: Keywords: hydrolase; serine proteinase
F: 46-87/Domain: fibronectin type II repeat homology <1F2>
F: 134-169/Domain: fibronectin type I repeat homology <FB1>
F: 177-208/Domain: EGF homology <EGF>
F: 216-294/Domain: kringle homology <KRG>
F: 359-597/Domain: trypsin homology <TRY>

Query Match 31.2%; Score 718.5; DB 2; Length 603;
Best Local Similarity 36.8%; Pred. No. 1e-47;
Matches 161; Conservative 64; Mismatches 149; Indels 63; Gaps 13;

Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 182 CLNGGRCLE---VEGHHLCDCPMGYTGFCDLDTTASCYEGRGVSYRGMAATTVSGAKCQ 238
Qy 73 PWSNATVLOQTHAHRSD-ALQLGLGKHNYCRNPDNRPRWCYVQVGLKPLVQBCMVHDC 131
Db 239 RWAS-----EATYRNWTAQALRGLGHTTCRNFNDNDRTPWCFVWMGNLSWEYCDLAQC 294
Qy 132 ADGKPSPEELKFO-----CGQKTLRPR 156
Db 295 QYPPQPTATPHD-RFEHPKLPSSRLSIQTPTQNALANELPETSSLLCGOR-LRKR 352
Qy 157 F---KIIGGEFTTIENQPFPAAYRRHRGSGSVTVCGGSLISPCWVISATHCFIDYPKK 212
Db 353 LSSLSRIVGLVULPAGHPYIAALY-----WGS--NFCSSGLIAPCWLTIAHCLQNRPAP 406
Qy 213 EDYTVILGRSLRNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCACQPSRTTQTI 271
Db 407 EELKVLGQDRHNSQCEHCQTAVHSYRLHEAFSPS--SYLNDLALLRLQKSDAGSCAQL 464
Qy 272 SRTQTCLPSMYNDPQFG--TSCEITGFGKNSTDYLPQELKMTVVKLISHRECCQPH 329
Db 465 SPYQTVCLPSGAPPSESETTCCVAGWGQHFEGAEYSFSLQEAQVPLISSRCSPE 524
Qy 330 YGSEVITKMLCAADPQWKTDCQSGSGGLVCSLQGRMTLTGIVSGRCALKDKPGVYTRVSHF 398
```


C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
F;75-106/Domain: EGF homology <EG1>
F;113-145/Domain: EGF homology <EG2>
F;152-185/Domain: EGF homology <EG3>
F;192-274/Domain: kringle homology <KRI>
F;312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS>
F;312-548/Domain: trypsin homology <TRY>

Query Match 29.3%; Score 674.5; DB 2; Length 558;
Best Local Similarity 38.0%; Pred. No. 2.3e-44;
Matches 159; Conservative 56; Mismatches 164; Indels 37; Gaps 16;
QY 13 CLNGGTCVSNKYENIHWNCNPKKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
DB 157 CQNGGVCSSRRHRSRF-TCACPDQYKGFCEIGPD-DCYVGDGYSGYRGKSVKTNQNPCL 214
QY 73 PWSNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYQVGLKPLVQE-CMVHDC 131
DB 215 YWNSHLLQETYNMFMEDAETHGIAEHNFRCNPDGDHKKPWCFCVKNSEKWKWEYCDYTV 274
QY 132 ADGKKPSPPPEEL-----KFO-CGOKTLRPRF--KIIGGEFTTIENQFWFAIY--- 177
DB 275 PVPDTP-NPVESLLEPVMELPFCFESCGKTEVAEHAHVRIYGGFKSTAGKHPWQVLSQTS 333
QY 178 -----RRHRGGSVTVVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMK 233
DB 334 PLATSMFQGG---HFCGALIHPCWLVTAHC-TDINTKHLKV-LGQDLKKTESHEQT 387
QY 234 FEVENLILHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQF--GT 291
DB 388 FRVEKIFKYSHYNERDEIPHNDIALKLPVGDGHCALESXVVKTVCLP---DGFPSGS 444
QY 293 CEITGFGKENSTLYPEQLKMTVVKLIHRECCQPHYGSSEVTTKMLCAADPQWK-TDS 351
DB 445 CHISGWSVTERGE--GSRQLLDAKVKLIANPLCNSRQLYDHTIDDSMICAGNLQKPGSDT 502
QY 352 CQGDGSGPLVCSLQGRWTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHKTEENG 409
DB 503 CQGDGSGPLTCEKDGTYGVYVGSWGQECG--KRPGVYTVQVTKFLNWKITMTMREAGL 558

RESULT 20
JC4795
plasma hyaluronan-binding protein precursor - human
N;Alternate names: hepatocyte growth factor activator-like protein; PHBP
N;Contains: serine proteinase (EC 3.4.21.-)
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
C;Accession: JC4795
R;Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
J. Biochem. 119, 1157-1165, 1996
A;Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP) r activator.
A;Reference number: JC4795; MUID:96425001; PMID:8827452
A;Accession: JC4795
A;Molecule type: mRNA
A;Residues: 1-560 <CHO>
A;Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
A;Experimental source: plasma
A;Note: parts of this sequence, including the amino ends of the mature chains, were determined.
C;Gene: GDB:HABP2; HABP; PHBP; HGFA1
A;Cross-references: GDB:4573962
C;Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; th
C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
C;Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-311/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
F;77-108/Domain: EGF homology <EG1>
F;115-147/Domain: EGF homology <EG2>
F;154-187/Domain: EGF homology <EG3>

F;194-276/Domain: kringle homology <KRI>
F;314-550/Domain: trypsin homology <TRY>
F;314-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
F;54,207/Binding site: carbohydrate (Asn)
F;77-88,82-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,24
F;362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 28.7%; Score 661.5; DB 1; Length 560;
Best Local Similarity 37.1%; Pred. No. 2.3e-43;
Matches 155; Conservative 58; Mismatches 166; Indels 39; Gaps 15;

QY 13 CLNGGTCVSNKYENIHWNCNPKKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
DB 159 CQNGATCSRHRRSKF-TCACPDQYKGFCEIG-SDDCYVGDGYSGYRGKSNRTVQNHACL 216
QY 73 PWSNATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYQVGLKPL-----VQEC 126
DB 217 YWNSHLLQETYNMFMEDAETHGIEHNFRCNPDADKPPCFIKVTNDKVKWEYCDVSAC 276
QY 127 MVHDCA-DGKKPSPPPEELK--FCCGQKTLRPR--FKIIGGEFTTIENQFWFAIYRRHR 181
DB 277 SAQDVAYPEESPTPESTKLPFGDSCGKTEAERIKRIYGGFKSTAGKHPWQASLQ---- 332
QY 182 GGSVT-----YVCGSLISPCWISATHCFIDYPKKEDYIVYLGSRSLNSNTQGMK 233
DB 333 -SSLPLTISMPQGHFCGALIHPCWLVTAHC-TDI-KTRHLKVVLGGDQLKKEEFHEQS 389
QY 234 FEVENLILHKDYSADTLAHNDIALKIRSEKGRCAQPSRTIQTICLPSMYNDPQF--GT 291
DB 390 FRVEKIFKYSHYNERDEIPHNDIALKLPVGDGHCALESXVVKTVCLP---DGFPSGS 445
QY 292 SCEITGFGKENSTLYPEQLKMTVVKLIHRECCQPHYGSSEVTTKMLCAADPQWK-TD 350
DB 446 ECHISGWSVTERGE--GSRQLLDAKVKLIANTLNSRQLYDHMIDDSMICAGNLQKPGQD 503
QY 351 SCQDGSGLVCSLQGRWTLTGIVSWRGCAKDKPGVYTRVSHFLPWIRSHKTEENG 408
DB 504 TCQDGSGLTCEKDGTYGVYVGSWGLEGCG--KRPGVYTVQVTKFLNWKITIKESG 559

RESULT 21
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N;Alternate names: Hageman factor (activated)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C;Accession: S45281; A61329
R;Shibuya, Y.; Samba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): cor
A;Reference number: S45281; MUID:94242782; PMID:8186251
A;Accession: S45281
A;Molecule type: mRNA
A;Residues: 1-593 <SHI>
A;Cross-references: GB:S70164
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
is, and ATC for residue 505 as Leu
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).
A;Reference number: A61329; MUID:77182112; PMID:861210
A;Accession: A61329
A;Molecule type: protein
A;Residues: 10-16; 'X', 18-19;525-550 <FUG>
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F;37-78/Domain: fibronectin type II repeat homology <1F2>
F;88-120/Domain: EGF homology <EGF>
F;125-160/Domain: fibronectin type I repeat homology <FB1>
F;207-287/Domain: kringle homology <KR>
F;350-587/Domain: trypsin homology <TRY>
F;541/Active site: Ser #status predicted

F;1-77/D	Domain:	activation peptide	#status predicted <APT>
F;78-560/P	Product:	plasmin chain A	#status predicted <ACH>
F;84-162/D	Domain:	kringle homology <KR1>	
F;166-243/D	Domain:	kringle homology <KR2>	
F;256-333/D	Domain:	kringle homology <KR3>	
F;358-435/D	Domain:	kringle homology <KR4>	
F;450-790/P	Product:	mini-plasminogen	#status experimental <MIN>
F;461-540/D	Domain:	kringle homology <KR5>	
F;561-790/P	Product:	plasmin chain B	#status experimental <BCH>
F;561-783/D	Domain:	trypsin homology <TRY>	
F;30-54,	34-42,84-162,103-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305-	bonds:	#status predicted
F;602,645,740/A	Active site:	His, Asp, Ser	#status predicted

Query Match	21.9%;	Score	504.5;	DB	1;	Length	790;	
Best Local Similarity	35.4%;	Pred.	No.	4.4e-31;	Mismatches	141;	Gaps	14;
Matches	130;	Conservative	45;	Indels	51;			

Qy	45	DKSKTCEYGNHFGHYRGKASTDTMGRCLPWN SATVLQOTVHAHR----	SDALQLGLGKHN	100
Db	456	DLSEDCWFGNGKYRGRKRAITVAGVPQENAA-----QP	RHSIFTPETNPRAGLEK-N	509
Qy	101	YCRNPD--NRRRPWCVVGKLPLVOECNHDCADGKKPSPPBELKPQCQCKTLRP---	156	
Db	510	YCRNPDGDNDGPWCYT-TNQKLFDPYCDVPCCVTS-----SFDCKPKVEPKKCP	558	
Qy	157	KFIIGBEFTTIENOPFAAIIYYRHRRGSVTYVCGSLISPCWVISATHC FIDYPKEDI	216	
Db	559	ARVVGGCIVPHSPFWQIISRYYRG----HFQGTILISPEWLTA KHCLEKSSSSSYK	614	
Qy	217	VYLGRSLNSNTQGEMKFVENLILHKDYSGADTLLAHNDIALLKIRSKRCQAQPSRTIQ	276	
Db	615	VILGAHEYHLGEGVQBIWSKLF--KEPS-----EADIALLKL SSP----	AVIDTKVI	662
Qy	277	TICLPSTMYNDPQF-----GTSCELTFGCXENSTDLYPEQLKM TVVKLI SHRECOQHYYG	332	
Db	663	PACLUPT-----PNYVADR TACYITGWGETKGT--YGAG LUKEARLP VIEKNVCNRVEYLG	716	
Qy	333	SEVTTMKLCAADPOWKTDSCQDGGSGGPLVCSLOGRWMTLTGI VSWGRCALKDKPGVVTRV	392	
Db	717	GKVSPNELCAHLAGGIDSQDGSGGPLVCFFKDXKIQLQ IGVTSWGLGCALPNKPGVVVRV	776	
Qy	393	SHELPLWI	399	
Db	777	SRFVTWI	783	

RESULT 24

PLHU

```

plasmin (EC 3.4.21.7) precursor [validated] - human
N/Alternate names: plasminogen precursor [misnomer]
N/Contains: angiotatin; microplasmin; plasminogen
C/Species: Homo sapiens (man)
C/Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 15-Sep-2000
C/Accession: A35229; I52242; A36646; I62738; I84609; S03735; A00929; A04627; A04625; A04626; A04627; A04628; A04629; A04630; A04631; A04632; A04633; A04634; A04635; A04636; A04637; A04638; A04639; A04640; A04641; A04642; A04643; A04644; A04645; A04646; A04647; A04648; A04649; A04650; A04651; A04652; A04653; A04654; A04655; A04656; A04657; A04658; A04659; A04660; A04661; A04662; A04663; A04664; A04665; A04666; A04667; A04668; A04669; A04670; A04671; A04672; A04673; A04674; A04675; A04676; A04677; A04678; A04679; A04680; A04681; A04682; A04683; A04684; A04685; A04686; A04687; A04688; A04689; A04690; A04691; A04692; A04693; A04694; A04695; A04696; A04697; A04698; A04699; A04700; A04701; A04702; A04703; A04704; A04705; A04706; A04707; A04708; A04709; A04710; A04711; A04712; A04713; A04714; A04715; A04716; A04717; A04718; A04719; A04720; A04721; A04722; A04723; A04724; A04725; A04726; A04727; A04728; A04729; A04730; A04731; A04732; A04733; A04734; A04735; A04736; A04737; A04738; A04739; A04740; A04741; A04742; A04743; A04744; A04745; A04746; A04747; A04748; A04749; A04750; A04751; A04752; A04753; A04754; A04755; A04756; A04757; A04758; A04759; A04760; A04761; A04762; A04763; A04764; A04765; A04766; A04767; A04768; A04769; A04770; A04771; A04772; A04773; A04774; A04775; A04776; A04777; A04778; A04779; A04780; A04781; A04782; A04783; A04784; A04785; A04786; A04787; A04788; A04789; A04790; A04791; A04792; A04793; A04794; A04795; A04796; A04797; A04798; A04799; A04800; A04801; A04802; A04803; A04804; A04805; A04806; A04807; A04808; A04809; A04810; A04811; A04812; A04813; A04814; A04815; A04816; A04817; A04818; A04819; A04820; A04821; A04822; A04823; A04824; A04825; A04826; A04827; A04828; A04829; A04830; A04831; A04832; A04833; A04834; A04835; A04836; A04837; A04838; A04839; A04840; A04841; A04842; A04843; A04844; A04845; A04846; A04847; A04848; A04849; A04850; A04851; A04852; A04853; A04854; A04855; A04856; A04857; A04858; A04859; A04860; A04861; A04862; A04863; A04864; A04865; A04866; A04867; A04868; A04869; A04870; A04871; A04872; A04873; A04874; A04875; A04876; A04877; A04878; A04879; A04880; A04881; A04882; A04883; A04884; A04885; A04886; A04887; A04888; A04889; A04890; A04891; A04892; A04893; A04894; A04895; A04896; A04897; A04898; A04899; A04900; A04901; A04902; A04903; A04904; A04905; A04906; A04907; A04908; A04909; A04910; A04911; A04912; A04913; A04914; A04915; A04916; A04917; A04918; A04919; A04920; A04921; A04922; A04923; A04924; A04925; A04926; A04927; A04928; A04929; A04930; A04931; A04932; A04933; A04934; A04935; A04936; A04937; A04938; A04939; A04940; A04941; A04942; A0
```

R; Forsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O. FEBS Lett. 213, 254-260, 1987

A:Title: Molecular cloning and characterization of a full-length cDNA clone for human I

A:Reference number: A26646; MUID:87162490; PMID:3030813

A:Accession: A26646

A:Molecule type: mRNA

A:Residues: 1-471, 'D', 473-810 <FOR>

A:Cross-references: GB:X05199; NID:g35530; PIDN:CAA28831.1; PID:g35531

A:Experimental source: liver

R; Malinowski, D.P.; Sadler, J.E.; Davie, E.W. Biochemistry 23, 4243-4250, 1984

A:Title: Characterization of a complementary deoxyribonucleic acid coding for human an

A:Reference number: 145961; MUID:85023311; PMID:6148961

A:Accession: 162738

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 232-471, 'D', 473-810 <MAL2>

A:Cross-references: GB:X02922; NID:g190112; PIDN:AAA60124.1; PID:g387031

A:Accession: 184609

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 367-419 <MAL3>

A:Cross-references: GB:X02921; NID:g190110; PIDN:AAA60123.1; PID:g190111

R; Brunscholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M. Eur. J. Biochem. 114, 465-470, 1981

A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human

A:Reference number: S03735; MUID:81212097; PMID:7238497

A:Accession: S03735

A:Molecule type: protein

A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>

R; Wiman, B. Eur. J. Biochem. 76, 129-137, 1977

A:Title: Primary structure of the B-chain of human plasmin.

A:Reference number: A04627; MUID:77225245; PMID:142009

A:Accession: A04627

A:Molecule type: protein

A:Residues: 581-810 <W1>

R; Wiman, B.; Wallen, P. Eur. J. Biochem. 58, 539-547, 1975

A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen t

A:Reference number: A04626; MUID:76043692; PMID:126863

A:Accession: A04626

A:Molecule type: protein

A:Residues: 483-507, 'E', 509-604 <W13>

R; Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L. J. Biol. Chem. 248, 1631-1633, 1973

A:Title: The primary structure of human plasminogen. II. The histidine loop of human p

A:Reference number: A92125; MUID:73149248; PMID:4694729

A:Contents: annotation; active site

R; Groskopf, W.R.; Summaria, L.; Robbins, K.C. J. Biol. Chem. 244, 3590-3597, 1969

A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of

A:Reference number: A92048; MUID:69234739; PMID:4240117

A:Contents: annotation; active site

R; Texter, M.; Vail, Z.; Pathy, L. J. Biol. Chem. 257, 7401-7406, 1982

A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminoge

A:Reference number: A92382; MUID:82213905; PMID:6919539

A:Contents: annotation; omega-aminocarboxylic acid binding sites

R.Vali, Z.; Pathy, L.
 J. Biol. Chem. 259, 13690-13694, 1984
 A;Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential
 A;Reference number: A92458; PMID:85054794; PMID:6094526
 A;Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site
 R;Cao, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Soehndel, S.; McCance, S.G.;
 J. Biol. Chem. 271, 29461-29467, 1996
 A;Title: Kringle domains of human angiotensin. Characterization of the anti-proliferative
 A;Reference number: A58811; PMID:97067211; PMID:8910613
 A;Contents: annotation
 R;Lijnen, H.R.; Ugwu, F.; Bini, A.; Collen, D.
 Biochemistry 37, 4699-4702, 1998
 A;Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (M)
 A;Reference number: A58812; PMID:9548733; PMID:9548733
 A;Contents: annotation
 R;Tullinsky, A.; Mulichak, A.M.
 submitted to the Brookhaven Protein Data Bank, July 1991
 A;Reference number: A51341; PDB:1PK4
 A;Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454
 R;Tullinsky, A.; Wu, T.P.
 submitted to the Brookhaven Protein Data Bank, August 1993
 A;Reference number: A51911; PDB:1PKR
 A;Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181
 R;Padmanabhan, K.; Tullinsky, A.
 submitted to the Brookhaven Protein Data Bank, April 1994
 A;Reference number: A52408; PDB:1PMK
 A;Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454
 R;Tullinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, December 1995
 A;Reference number: A65244; PDB:1CEA
 A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
 R;Tullinsky, A.; Mathews, I.I.
 submitted to the Brookhaven Protein Data Bank, December 1995
 A;Reference number: A65245; PDB:1CEB
 A;Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181
 R;Mulichak, A.M.; Tullinsky, A.; Ravichandran, K.G.
 Biochemistry 30, 10576-10588, 1991
 A;Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å
 A;Reference number: A58819; PMID:92031502; PMID:1657148
 A;Contents: annotation
 R;Wu, T.P.; Padmanabhan, K.; Tullinsky, A.; Mulichak, A.M.
 Biochemistry 30, 10589-10594, 1991
 A;Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin
 A;Reference number: A58818; PMID:92031503; PMID:1657149
 A;Contents: annotation
 R;De Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tullinsky, A.; Westbrook, M.
 Biochemistry 31, 270-279, 1992
 A;Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.1
 A;Reference number: A39483; PMID:92118803; PMID:1310033
 A;Contents: annotation; X-ray crystallography, 2.4 angstroms
 R;Stec, B.; Teeter, M.M.; Whittow, M.; Yamano, A.
 submitted to the Brookhaven Protein Data Bank, June 1995
 A;Reference number: A65980; PDB:1KRN
 A;Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454
 R;Rejante, M.; Llinas, M.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A;Reference number: A65803; PDB:1HPJ
 A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
 R;Rejante, M.; Llinas, M.
 submitted to the Brookhaven Protein Data Bank, August 1996
 A;Reference number: A65804; PDB:1HPK
 A;Contents: annotation; conformation by (1)H-NMR, residues 103-181
 R;Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 927-937, 1994
 A;Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.
 A;Reference number: S43645; PMID:94237157; PMID:8181475
 A;Contents: annotation; conformation by (1)H-NMR, residues 96-184
 R;Rejante, M.R.; Llinas, M.
 Eur. J. Biochem. 221, 939-949, 1994

A;Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasmin
 A;Reference number: A58817; PMID:94237158; PMID:8181476
 A;Contents: annotation; conformation by (1)H-NMR
 C;Comment: Plasminogen is synthesized by the kidney and is present in plasma and many c
 d PIR:FGHUB).
 C;Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately a
 rg-580, resulting in two chains connected by two disulfide bonds. Without the inhibitor
 C;Comment: Microplasmin is formed by autolytic cleavage of plasmin under artificial co
 C;Comment: Stromelysin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. 1
 ting solid tumors.
 C;Genetics:
 A;Gene: GDB:PLG
 A;Cross-references: GDB:119498; OMIM:173350
 A;Map position: 6q26-6q27
 A;Introns: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 52
 C;Function:
 A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a v
 us the walls of the graafian follicle; also activates the urokinase-type plasminogen ac
 C;Pathway: fibrinolysis
 C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolog
 C;Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyc
 F;1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F;1-19/Domain: signal sequence #status predicted <SIG>
 F;20-810/Product: plasminogen #status experimental <PRO>
 F;20-96/Domain: activation peptide #status experimental <APT>
 F;79-466/Product: angiotensin #status experimental <AST>
 F;97-580.581-810/Product: plasmin #status experimental <MAT>
 F;97-580/Domain: plasmin chain A #status experimental <CHA>
 F;103-181/Domain: kringle homology <KR1>
 F;185-262/Domain: kringle homology <KR2>
 F;275-352/Domain: kringle homology <KR3>
 F;377-454/Domain: kringle homology <KR4>
 F;481-560/Domain: kringle homology <KR5>
 F;550-580.581-810/Product: microplasmin #status experimental <MMT>
 Query Match 21.8%; Score 501; DB 1; Length 810;
 Best Local Similarity 33.5%; Pred. No. 8.3e-31;
 Matches 139; Conservative 45; Mismatches 167; Indels 64; Gaps 16;
 QY 10 NCDCLNGGTCVKNKYFNSIHWNCNPKKFGQ-----HCEIDKSKTCYEGNGH 56
 DB 428 NPDADKGPWCFTTDSVRWEYNLKKCSGTASVAPPVLLPNVETPSEEDCMFGNGK 487
 QY 57 FYRGKASTDMGRPCLPWNSATVLOQTYHAHR----SDALQLGLGKHNYCRNPD-NRRRP 111
 DB 488 GYRGKRAITVTGTPCODMAA-----QEPHRHSIFTPTNPRAGLEK-NYCRNPDGVDGVP 541
 QY 112 WCVYVQVLKPLVQECMVHDCADGKXSPPEBLKFCQCGQKTLRPR---PKIIGGEFTTIE 168
 DB 542 WCYT-TNPKRLDYCDVPOCA---APS-----FDCGRQPVEPKKCPGRVVGCVAPH 590
 QY 169 NOPWFAAIYRRHRGSGVTVCVCGSLISPCOWISATHCFIDYPKEDYIVVIGRSLNSNT 228
 DB 591 SWPQVSL--RTRFG--MHFCGGLISPEWVTAHCKLEKSPRPSYKVLGAHQEWLIE 646
 QY 229 QEMKFEVENLLHKDYSADTLAHNDIALLKIRSKGRCQAQPSRTIQTICLPSMYNDPQ 288
 DB 647 PHVQIEVSRLEPTEPT-----RKDIALLKLSPP-----AVITDKVIPACLPS----FN 690
 QY 289 F----GTSCEITGFGKENSTDYLYPEQLKMTVVVLLKISHRECOQPHYSSEVTKMLCAAD 344
 DB 691 YVADRTCEFTGWTGTEGT--FGAGLKEAQLPVLENKVCNRYEFLNGRVSOSTELCAGH 748
 QY 345 POWKTDSCQSGSGGLVCSLQGRMTLTGIVSWGRCALKDKPGVTVTRVSHFLPWI 399
 DB 749 LAGGTDSCQSGSGGLVCFEKDKYILQGVTSWGLGCARPKNKPGVTVRVSFRFTWI 803
 RESULT 25
 B30848
 plasmin (EC 3.4.21.7) precursor - rhesus macaque
 C;Species: Macaca mulatta (rhesus macaque)

C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 22-Jun-1999
 C/Accession: B32869; B30848
 R/Tomlinson, J.E.; McLean, J.W.; Law, R.M.
 J. Biol. Chem. 264, 5957-5965, 1989
 A/Title: Rhesus monkey apolipoprotein(a), Sequence, evolution, and sites of synthesis.
 A/Reference number: A32869; MUID:89174660; PMID:2925643
 A/Accession: B32869
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-810 <TOM>
 A/Cross-references: GB:J04697; NID:G342272; PID:AAA36901.1; PID:G342273
 C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C/Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>
 F:1-96/Domain: signal sequence #status predicted <SIG>
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:377-454/Domain: kringle homology <KR4>
 F:481-560/Domain: kringle homology <KR5>
 F:581-803/Domain: trypsin homology <TRY>
 F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 324-352, 365-760/Active site: His, Asp, Ser #status predicted
 bonds: #status predicted
 F:622, 665, 760/Active site: His, Asp, Ser #status predicted

Query Match 21.6%; Score 497; DB 2; Length 810;
 Best Local Similarity 33.7%; Pred. No. 1.7e-30;
 Matches 140; Conservative 46; Mismatches 165; Indels 64; Gaps 16;

QY 10 NCCLNGGTCVSNKYFSNIHWCNCPKFGGQ-----HCEIDSKTCYEGNGH 56
 Db 428 NPADKGPWCFTTDPSPVRWEYCNLKKCSGTEGSAAPPVPAQLPDAETPSEEDCMFNGK 487

QY 57 FYRGKASTDTMGPRCLPWNATVLOQTYHAH-----SDALQLGLGKHNYCRNPD-NRRRP 111
 Db 488 GYRGKATVTAAGTFCQGWAA-----QEPHRSIFTPTNPRADL-----EKNYCRNPDG 536

QY 112 WCYVQVGLKPLVQECMHDCADGKSPPEELKFCQCKTLRPR---FKIIGFEFTIE 168
 Db 542 WCYT-TNPRKLFYCDVPQCAAS-----SFDCKGKQVPPKPCFGRVVGCVAYPH 590

QY 169 NQWFAAIYRRHGGSVTVYVCGSLISPCWISATHCFIDYPKKEDIVYLGSRSLNSNT 228
 Db 591 SWPN-QISLRTRLG--MHFCGGTLISPEWVLAHCKLSRSPSFYKVLGAHR-----641

QY 229 QGEMKFEVENILHKDYSADTLAHNDIALKRSKEGRCAQPSRTIQTCLPSMNDPQ 288
 Db 642 EVHLEPHVQIEYSKMFSEPARA---DIALKLSPP-----AIITDKVIPACLPSPN 690

QY 289 F----GTSCETITGFKENSTDYLYPEOLKMTVVKLISHRECOQPHYYSVTTKMLCAAD 344
 Db 691 YVADRTECITGWGETQGT--YGAGLLKEARLPVIEKNVCNRYEFLNGTVKTELCAH 748

QY 345 POWKTDSCQDGGGLVCSLQGRMTLGTIVSWGRGCALDKPKGVYTVRSHFLPWI 399
 Db 749 LAGTDSQCQDGGGLVCFPEKDYILQGVTSWGLGCARPNKPGVYVRSRFTWI 803

RESULT 26
 PLMS
 N/plasmin (EC 3.4.21.7) precursor - mouse
 N/Contains: angiotensin; plasminogen
 C/Species: Mus musculus (house mouse)
 C/Date: 20-Sep-1991 #sequence_revision 01-Nov-1996 #text_change 18-Jun-1999
 C/Accession: A38514; S48202; S48203
 R/Degen, S.J.F.; Bell, S.M.; Schaefer, L.A.; Elliott, R.W.
 Genomics 8, 49-61, 1990
 A/Title: Characterization of the cDNA coding for mouse plasminogen and localization of
 A/Reference number: A38514; MUID:91184812; PMID:2081600
 A/Accession: A38514
 A/Molecule type: mRNA
 A/Residues: 1-812 <DEG>
 A/Cross-references: GB:J04766; NID:g200402; PID:AAA50168.1; PID:g200403

R/Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
 Eur. J. Biochem. 224, 863-871, 1994
 A/Title: Characterization of the murine plasma fibrinolytic system.
 A/Reference number: S48202; MUID:95010076; PMID:7523120
 A/Accession: S48202
 A/Molecule type: protein
 A/Residues: 20-25 <LIJ>
 A/Accession: S48203
 A/Molecule type: protein
 A/Residues: 22-27 <LIJ>
 C/Comment: Plasminogen is synthesized by the kidney and is present in plasma and many c
 immediately after dissociation from the clot. In the presence of the inhibitor, the activ
 e inhibitor, the activation involves also removal of the activation peptide.
 C/Comment: Stromelysin 1 (see PIR:KCMSS1) acts on plasminogen to produce angiotatin. 1
 eul in treating solid tumors.
 C/Function:
 A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a v
 ns the walls of the graafian follicle; also activates the urokinase-type plasminogen ac
 C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
 C/Keywords: angiotensin inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyc
 F:1-19/Domain: plasminogen-related protein precursor homology <PLPH>
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-812/Domain: plasminogen #status predicted <PRO>
 F:20-96/Domain: activation peptide #status predicted <APT>
 F:79-466/Domain: angiotensin #status predicted <AST>
 F:97-581, 582-812/Domain: plasmin #status predicted <MAT>
 F:97-581/Domain: chain A #status predicted <ACH>
 F:103-181/Domain: kringle homology <KR1>
 F:185-262/Domain: kringle homology <KR2>
 F:275-352/Domain: kringle homology <KR3>
 F:377-454/Domain: kringle homology <KR4>
 F:481-560/Domain: kringle homology <KR5>
 F:582-812/Domain: chain B #status predicted <BCH>
 F:582-805/Domain: trypsin homology <TRY>
 F:49-73, 53-61, 103-181, 124-164, 152-176, 185-262, 188-316, 206-245, 234-257, 275-352, 296-335, 324-352, 365-760/Active site: His, Asp, Ser #status predicted
 bonds: #status predicted
 F:78-79/Cleavage site: Glu-Asn (stromelysin 1) #status predicted
 F:136, 308/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:466-467/Cleavage site: Thr-Val (stromelysin 1) #status predicted
 F:581-582/Cleavage site: Arg-Val (plasminogen activator) #status experimental
 F:624, 667, 762/Active site: His, Asp, Ser #status predicted

Query Match 21.6%; Score 497; DB 1; Length 812;
 Best Local Similarity 31.7%; Pred. No. 1.7e-30;
 Matches 133; Conservative 50; Mismatches 165; Indels 72; Gaps 15;

QY 10 NCCLNGGTCVSNKYFSNIHWCNCPK--KFGQHCET-----DKSKTCYEGNGH 56
 Db 428 NPADKGPWCFTTDPSPVRWEYCNLKKCSGTEGSAAPPVPAQLPDAETPSEEDCMFNGK 487

QY 57 FYRGKASTDTMGPRCLPWNATVLOQTYHAH-----SDALQLGLGKHNYCRNPD- 106
 Db 488 DYRGKATVTAAGTFCQGWAA-----QEPHRSIFTPTNPRADL-----EKNYCRNPDG 536

QY 107 NRRRPWCYVQVGLKPLVQECMHDCADGKSPPEELKFCQCKTLRPR---FKIIGGE 163
 Db 537 DVNGPWCYT-TNPRKLYDYCDIPLCASAS-----SFECKGKQVPPKPCFGRVVGVC 586

QY 164 FTTIENQWFAAIYRRHGGSVTVYVCGSLISPCWISATHCFIDYPKKEDIVYLGSR 223
 Db 587 VANPHSWPQISLRTRFTG---QHFCGGTLISPEWVLAHCKLSRSPSFYKVLGAHE 643

QY 224 LNSNTQGMKFEVENILHKDYSADTLAHNDIALKRSKEGRCAQPSRTIQTCLPSM 283
 Db 644 EYIRGLDVQIEISVAKLLE-----PNRDIALLKL---SRPATITDKVIPACLPSPN 690

QY 284 YNDPQF----GTSCETITGFKENSTDYLYPEOLKMTVVKLISHRECOQPHYYSVTTK 339
 Db 691 ---PNTWVADRTECITGWGETQGT--FGAGRLKEARLPVIEKNVCNRYEFLNRRVKSTE 745

QY 340 LCAADPQWKTDSQCQDGGGLVCSLQGRMTLGTIVSWGRGCALDKPKGVYTVRSHFLPWI 399

QY 265 EGRCAQPSRTIQICLIPSMYNDQF-----GTSCEITGFGKSTDYLYPQLKMTVVKLI 320
Db 1282 -SRPAITDKVIPACLPS-----PNYVITAMTECYITGWTGQT--FGAGLLKEAQLHVI 1334
QY 321 SHRECCQPHYGVGVTTKMLCAADPQWKTDSCQDSGGPLVCSLQGRMTLTGVSVGRGC 380
Db 1335 ENTVCNHYEFLNGRVKSTELCAGHLAGGTDRCCQDNGGPPVCPDKXILRGITSWGPGC 1394
QY 381 ALKDKPGVYTRVSHFLDWI 399
Db 1395 ACNPKPGVYRVVSFTWI 1413
RESULT 29
S00657
apoprotein(a) (EC 3.4.21.-) precursor [validated] - human
N/Alternate names: apolipoprotein(a); lipoprotein(a) chain apo(a)
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text change 08-Dec-2000
C/Accession: S00657; A28017; A47277; I60906; A47233; I52415; I65286
R/McLean, J.W.; Tomlinson, J.E.; Kuang, W.J.; Eaton, D.L.; Chen, E.Y.; Fless, G.M.; Scar
Nature 330, 132-137, 1987
A/Title: cDNA sequence of human apolipoprotein(a) is homologous to plasminogen.
A/Reference number: S00657; MUID:88039109; PMID:3670400
A/Accession: S00657
A/Molecule type: mRNA
A/Residues: 1-4548 <MCL>
A/Cross-references: GB:X06290; EMBL:X06696; NID:g28619; PIDN:CAA29618.1; PID:g28620
R/Eaton, D.L.; Fless, G.M.; Kohr, W.J.; McLean, J.W.; Xu, Q.T.; Miller, C.G.; Lawn, R.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 3224-3228, 1987
A/Title: Partial amino acid sequence of apolipoprotein(a) shows that it is homologous to
A/Reference number: A28017; MUID:87204109; PMID:3472206
A/Accession: A28017
A/Molecule type: protein
A/Residues: 20-21, 'P', 23-34, 177-179, 'N', 181-186, 'T', 188-196, 'DKG', 200:292-314, 'W', 316-31
X', 4396-4401 <EAT>
R/Wade, D.P.; Clarke, J.G.; Lindahl, G.E.; Liu, A.C.; Zysow, B.R.; Meer, K.; Schwartz, K
Proc. Natl. Acad. Sci. U.S.A. 90, 1369-1373, 1993
A/Title: 5' control regions of the apolipoprotein(a) gene and members of the related pla
A/Reference number: A47277; MUID:93165698; PMID:7679504
A/Accession: A47277
A/Status: preliminary; translation not shown; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <RES>
A/Cross-references: GB:I07899; NID:g967973; PID:g967974
R/Margaret, N.; Acquati, F.; Magnaghi, P.; Bruno, L.; Pontoglio, M.; Rocchi, M.; Sacco
Proc. Natl. Acad. Sci. U.S.A. 89, 11584-11588, 1992
A/Title: Characterization by yeast artificial chromosome cloning of the linked apolipop
A/Reference number: A47233; MUID:93087573; PMID:1454851
A/Accession: I60906
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <RE2>
A/Cross-references: GB:M90078; NID:g178786; PIDN:AAA35547.1; PID:g553188
A/Note: apo(a) gene 1 (nomenclature of reference I52415)
A/Accession: A47233
A/Status: preliminary; translation not shown; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <RE5>
A/Cross-references: GB:M90079; NID:g178784; PIDN:AAA35546.1; PID:g553187
R/ichinose, A.
Biochemistry 31, 3113-3118, 1992
A/Title: Multiple members of the plasminogen-apolipoprotein(a) gene family associated wi
A/Reference number: I52415; MUID:92207924; PMID:1554698
A/Accession: I52415
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-16 <RE3>
A/Cross-references: GB:M86877; NID:g178780; PIDN:AAB49909.1; PID:g553185
A/Note: apo(a) gene 1 (nomenclature of reference I52415)
A/Accession: I65286
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA

A/Residues: 1-16 <RE4>
A/Cross-references: GB:M86878; NID:g178782; PIDN:AAA51749.1; PID:g553186
C/Genetics:
A/Gene: GDB:LPA
A/Cross-references: GDB:120699; OMIM:152200
A/Map position: 6q26-6q27
A/Note: several genes closely linked on chromosome 6 are identical in the first coding
rs of kringle repeats
C/Superfamily: apolipoprotein(a); kringle homology; trypsin homology
C/Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-4548/Product: apolipoprotein(a) #status experimental <MAT>
F/28-105/Domain: kringle homology <KR1>
F/142-219/Domain: kringle homology <KR2>
F/256-333/Domain: kringle homology <KR3>
F/370-447/Domain: kringle homology <KR4>
F/484-561/Domain: kringle homology <KR5>
F/598-675/Domain: kringle homology <KR6>
F/712-789/Domain: kringle homology <KR7>
F/826-903/Domain: kringle homology <KR8>
F/940-1017/Domain: kringle homology <KR9>
F/1054-1131/Domain: kringle homology <KR10>
F/1168-1245/Domain: kringle homology <KR11>
F/1282-1359/Domain: kringle homology <KR12>
F/1396-1473/Domain: kringle homology <KR13>
F/1510-1587/Domain: kringle homology <KR14>
F/1624-1701/Domain: kringle homology <KR15>
F/1738-1815/Domain: kringle homology <KR16>
F/1852-1929/Domain: kringle homology <KR17>
F/1966-2043/Domain: kringle homology <KR18>
F/2080-2157/Domain: kringle homology <KR19>
F/2194-2271/Domain: kringle homology <KR20>
F/2308-2385/Domain: kringle homology <KR21>
F/2422-2499/Domain: kringle homology <KR22>
F/2536-2613/Domain: kringle homology <KR23>
F/2650-2727/Domain: kringle homology <KR24>
F/2764-2841/Domain: kringle homology <KR25>
F/2878-2955/Domain: kringle homology <KR26>
F/2992-3069/Domain: kringle homology <KR27>
F/3106-3183/Domain: kringle homology <KR28>
F/3220-3297/Domain: kringle homology <KR29>
F/3334-3411/Domain: kringle homology <KR30>
F/3448-3525/Domain: kringle homology <KR31>
F/3562-3639/Domain: kringle homology <KR32>
F/3676-3753/Domain: kringle homology <KR33>
F/3782-3859/Domain: kringle homology <KR34>
F/3896-3973/Domain: kringle homology <KR35>
F/4010-4087/Domain: kringle homology <KR36>
F/4124-4201/Domain: kringle homology <KR37>
F/4228-4307/Domain: kringle homology <KR38>
F/4328-4541/Domain: trypsin homology <TRY>
Query Match 20.6%; Score 474; DB 1; Length 4548;
Best Local Similarity 35.2%; Pred. No. 6.5e-28;
Matches 129; Conservative 36; Mismatches 133; Indels 68; Gaps 17;
QY 50 CYEGNGHFRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGX-----HNYCRN 104
Db 4228 CMFGNGKGYRGKKAATVTTGTCQEW-----AAQEPHRRHSTFIPGNTKWAGLEKNYCRN 4280
QY 105 PD-NRRRRPCWYQVGLKPLVQECMVHDCADGKKPSSPPEELKFCQCKTLRPR---FKLI 160
Db 4281 PGDINGPWCYT-MNPKLFDYCDIPLCASS-----SPDCKGPQVEPKKCPGSIV 4329
QY 161 GGEFTTIENQWPFAAIYRRHRRGVSVTYCGGSLISPCWVISATHCFTIDYPKKEDYIVILG 220
Db 4330 GGCVAHPHSPWQVSL--RTRFGK--HFCGGTILSEFWLTAHCLKKSRRSSYKVIILG 4385
QY 221 RRR---LNNTQGEKMFVENILHKDYSADTLAHNHDIALLKIRSKGRCQAPSTIOT 277
Db 4386 AHQEVNLESHVQ---EIEVSRFLFLEPTQA-----DIALKL-----SRPAVITDKVMP 4430
QY 278 ICLPSMYNDQF-----GTSCEITGFGKSTDYLYPQLKMTVVKLIHRECCQPHYYS 333

```
Db 4431 ACLPS-----PYMYTARTECYITNGEFGQT--FGTGLLKEAQLLVINEVCN--HY--- 4479
Qy 334 EVTTKMLCAADPQWKTDSQQDSGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVS 393
Db 4480 ----KYTCAEHLARGTDSQQDSGGPLVCFEKKVKYILQGVTSWGLGCARPKNKPGVYARVS 4535
Qy 394 HFLPMI 399
Db 4536 RFVTWI 4541
```

RESULT 30

```
A61545
Plasmin (EC 3.4.21.7) precursor - horse (fragments)
N:Alternate names: plasminogen
N:Contains: miniplasminogen
C:Species: Equus caballus (domestic horse)
C:Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C:Accession: A61545; S17527
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: A61545
A:Molecule type: protein
A:Residues: 1-33,34-117 <SCH>
R:Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A:Title: Complete amino acid sequence of equine miniplasminogen.
A:Reference number: S17527; MUID:92052077; PMID:1946332
A:Accession: S17527
A:Molecule type: protein
A:Residues: 118-455 <SC2>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase; z
F:1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
F:1-33/Domain: activation peptide (fragment) #status experimental <APT>
F:34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
F:37-114/Domain: kringle homology <KR4>
F:118-455/Product: miniplasminogen #status experimental <MIN>
F:126-205/Domain: kringle homology <KR5>
F:226-455/Domain: plasmin chain B #status experimental <BCH>
F:226-448/Domain: trypsin homology <TRY>
F:267,310,405/Active site: His, Asp, Ser #status predicted
```

```
Query Match 19.9%; Score 457.5; DB 2; Length 455;
Best Local Similarity 29.5%; Pred. No. 1e-27;
Matches 132; Conservative 45; Mismatches 156; Indels 115; Gaps 15;
```

```
Qy 40 QHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPMSATV--LQTTYAHRSDALQLGLG 97
Db 28 BECEA-KVQDCYQDKGSEYRGTSITVTGKKCQSWSNTPHHQKTPKYPNADLTW--- 83
Qy 98 KHNCRNPDNRPRWCYVQ-----VGLK---PLVQECMVHDC-----DGKK--- 136
Db 84 --NYCRNPDGDKGWCYTTDPSVWERFCNLKCKSETVQEPSEPDCLGIGKYGQKKATT 141
Qy 137 -----PSSPP----- 141
Db 142 VTGTRQAAQAQBPHRHSIFTPKANPWANLEKNYCRNPDGDVNGPWCYTMNPQKLFDYCD 201
Qy 142 ----EELKFOGQKTLRPR---FKIIGGEFTTIENQWFAALYRRHRGGSVTVVCGGSLI 194
Db 202 VPQCESSPFDGKPKVFKKSGRIVGGCVIAHSWV--QISLRTFRG--RHPCGGTLI 257
Qy 195 SPQWISATCFIDYPKKEDYIVYLGSRSLNSNTQGEKMFVENLILHKDYSADTLAHN 254
Db 258 SPEWVLTAAHCLERSRPSTYKVVLGTHHELRLAAGAAQIDVSKLFLEPSRA----- 309
Qy 255 DIALLKIRSKRCAQPSRIQTI---CLPSWYNDPOFGTSCEITGPKGENSTDLYPEQ 311
Db 310 DIALLKLS-----PAITTONVIPACLPADPPVYVANWAECFVTGWGE--TQDSSNAGV 360
```

```
Qy 312 LKMTVVKLISHRECCQPHYGYSEVTTTMLCAADPQWKTDSQQDSGGPLVCSLQGRMTLT 371
Db 361 LKEAQLPVIENKVCNRXEYLNGRYKSTELCAGHLVGGVDSQQDSGGPLVCFEKKYILQ 420
Qy 372 GIVSWGRCALKDKPGVYTRVSHFLPMI 399
Db 421 GVTWGLGCARPKNKPGVYTRVSSFINWI 448
```

Search completed: December 3, 2003, 14:44:15
Job time : 24.2586 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 13.2061 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNELHQVPSNCDLNGTGV.....VSHFLPWIRSHYKENGAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	length	DB ID	Description
1	2298	99.9	431	1 UROK HUMAN	P00749 homo sapien
2	2148	93.4	433	1 UROK PAPCY	P16227 papio cynoc
3	1859.5	80.8	442	1 UROK PIG	P04185 sus scrofa
4	1767	76.8	433	1 UROK BOVIN	Q05589 bos taurus
5	1690.5	73.5	432	1 UROK RAT	P29598 rattus norv
6	1660.5	72.2	433	1 UROK MOUSE	P06869 mus musculu
7	1004	43.6	434	1 UROK CHICK	P15120 gallus gall
8	864.5	37.6	477	1 URT2 DESRO	P15638 desmodus ro
9	864.5	37.6	562	1 TPA HUMAN	P00750 homo sapien
10	858.5	37.3	431	1 URT8 DESRO	P98121 desmodus ro
11	844.5	36.7	559	1 TPA_EAT	P19637 rattus norv
12	836.5	36.4	477	1 URT1 DESRO	P98119 desmodus ro
13	826.5	35.9	559	1 TPA_MOUSE	P11214 mus musculu
14	818	35.5	566	1 TPA_BOVIN	Q28198 bos taurus
15	752	32.7	394	1 URTG DESRO	P49150 desmodus ro
16	735.5	32.0	655	1 HGFA HUMAN	Q04756 homo sapien
17	726	31.6	653	1 HGFA_MOUSE	Q9r098 mus musculu
18	718.5	31.2	603	1 FA12 CAVPO	Q04962 cavia porce
19	692	30.1	615	1 FA12 HUMAN	P00748 homo sapien
20	642	27.9	593	1 FA12 BOVIN	P98140 bos taurus
21	504.5	21.9	790	1 PLMN_PIG	P06867 sus scrofa
22	501	21.8	812	1 PLMN_MOUSE	P20918 mus musculu
23	499	21.7	810	1 PLMN_HUMAN	P00747 homo sapien
24	497	21.6	810	1 PLMN_MACMU	P12545 macaca mula
25	494.5	21.5	333	1 PLMN_CANFA	P80009 canis fami
26	493	21.4	343	1 PLMN_SHEEP	P81286 ovis aries
27	484.5	21.1	812	1 PLMN_BOVIN	P06868 bos taurus
28	484	21.0	1420	1 APOA_MACMU	P14417 macaca mula
29	474	20.6	4548	1 APOA_HUMAN	P08519 homo sapien
30	462	20.1	875	1 NETR_HUMAN	P56730 homo sapien
31	449.5	19.5	338	1 PLMN_HORSE	P80010 equus cabal
32	445	19.3	761	1 NETR_MOUSE	O08762 mus musculu
33	435	18.9	810	1 PLMN_ERIEU	Q29485 erinaceus e

RESULT 1

UROK_HUMAN STANDARD; PRT; 431 AA.

ID UROK_HUMAN STANDARD; PRT; 431 AA.
AC P00749; Q15844; Q16618; Q969W6;
DT 21-JUL-1986 (Rel. 01, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAUI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=85215647; PubMed=2987867;
RA Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;
RT "The human urokinase-plasminogen activator gene and its promoter.";
RL Nucleic Acids Res. 13:2759-2771(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,
RT Steffens G.J., Heyneker H.L.;
RT "Cloning and expression of the gene for pro-urokinase in Escherichia coli.";
RL Biotechnology 3:923-929(1985).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=86056954; PubMed=2415429;
RA Nagai M., Hiratsuka R., Kaneda T., Hayasuke N., Arimura H.,
RA Nishida M., Suyama T.;
RT "Molecular cloning of cDNA coding for human prepro-urokinase.";
RL Gene 36:183-188(1985).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=85203359; PubMed=3898571;
RA Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,
RA van Elsen A., Herzog A., Bollen A.;
RT "Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.";
RL DNA 4:139-146(1985).
RN [5]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA TISSUE=Lung;
RC MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

ALIGNMENTS

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[7]
RN SEQUENCE OF 66-431 FROM N.A.
RP MEDLINE=84272706; PubMed=6589620;
RX Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
RA "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
[8]
RN SEQUENCE OF 21-177.
RP MEDLINE=83055084; PubMed=6754569;
RX Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RA Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
[9]
RN SEQUENCE OF 156-176 AND 179-224.
RP MEDLINE=83003608; PubMed=6749491;
RX Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
RA Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
[10]
RN SEQUENCE OF 158-410.
RP MEDLINE=83055099; PubMed=6754572;
RX Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
RA "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
[11]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RP MEDLINE=96000858; PubMed=8591045;
RX Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RT urokinase-type plasminogen activator.";
RL Structure 3:681-691(1995).
[12]
RN X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RP MEDLINE=20266327; PubMed=10805774;
RX Sperl S., Jacob U., Arroyo de Prada N., Sturzbecher J., Wilhelm O.G.,
RA Boes W., Magdolen V., Huber R., Moroder L.;
RT "4-aminomethylphenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
[13]
RN STRUCTURE BY NMR.
RP MEDLINE=89127526; PubMed=2536903;
RX Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RA "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
[14]
RN STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
RL Biochemistry 31:9562-9571(1992).
[15]
RN STRUCTURE BY NMR OF 67-155.
RP MEDLINE=94149701; PubMed=8107091;
RX Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RA "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
RL J. Mol. Biol. 235:1548-1559(1994).
[16]
RN VARIANT LEU-141.
RP MEDLINE=96186279; PubMed=8652631;
RX Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
RA Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle
RT structure.";
RL Biochim. Biophys. Acta 1293:83-89(1996).
[17]
RN VARIANT LEU-141.
RP MEDLINE=97218551; PubMed=9065988;
RX Conne B., Berczy M., Belin D.;
RA "Detection of polymorphisms in the human urokinase-type plasminogen
RT activator gene.";
RL Thromb. Haemost. 77:434-435(1997).
[18]
RN VARIANT LEU-141.
RP MEDLINE=97337920; PubMed=9194591;
RX Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Heil W.,
RA Creutzburg S., Graeff H., Magdolen V.;
RT "Mutational analysis of the genes encoding urokinase-type plasminogen
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
RL Electrophoresis 18:686-689(1997).
CC -1- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
CC THERAPY OF THROMBOLYTIC DISORDERS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
CC -1- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02419; CAA26268.1; -
DR EMBL; M15476; AAA61253.1; -
DR EMBL; D00244; BAA00175.1; -
DR EMBL; D11143; BAA01919.1; -
DR EMBL; X02760; CAA26535.1; -
DR EMBL; AF377330; AAK53822.1; -
DR EMBL; BC013575; AAH13575.1; -
DR EMBL; K03226; AAC97138.1; -
DR EMBL; K02286; AAA61252.1; -
DR EMBL; A21571; CAA01559.1; -
DR EMBL; A18397; CAA01390.1; -
DR PIR; A00931; UKHU.
DR PDB; 1KDU; 31-OCT-93.


```
Query Match      99.9%; Score 2298; DB 1; Length 431;
Best Local Similarity 99.8%; Pred. No. 1.7e-174;
Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCKPFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCKPFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASDTMTGRCLPWNSTATVLTQYTHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 120
DB 81 KASDTMTGRCLPWNSTATVLTQYTHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 140
QY 121 PLVQECMVHDCADKPKSSPEELKFCQGGKTLRPRFKIIGBEFTTIENQPFWFAAIYRRH 180
DB 141 PLVQECMVHDCADKPKSSPEELKFCQGGKTLRPRFKIIGBEFTTIENQPFWFAAIYRRH 200
QY 181 RGSVTVVCGSLSPCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 240
DB 201 RGSVTVVCGSLSPCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 260
QY 241 LHKDYSADTLAHNDIALALKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCBITGFGK 300
DB 261 LHKDYSADTLAHNDIALALKIRSKRGCAQPSRTIQTICLPSMYNDPFGTSCBITGFGK 320
QY 301 ENSTDYLYPQLKMTVVKLSHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 360
DB 321 ENSTDYLYPQLKMTVVKLSHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSGRGKALDKPKGVYTVSHFLPWIRSHTKBEENGLAL 411
DB 381 VCSLQGRMTLTGIVSGRGKALDKPKGVYTVSHFLPWIRSHTKBEENGLAL 431

RESULT 2
UROK_PAPCY
ID_UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PAU.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RT plasminogen activator.";
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/)
CC -----

or send an email to license@isb-sib.ch.
CC -----
CC EMBL; X51935; CAA36200.1; -.
CC PIR; S14687; UKBAY.
CC HSP; P00749; ILMW.
CC MEROPS; S01.231; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser.protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS50240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
CC Kringle; EGF-like domain; Zymogen; Signal.
CC SIGNAL 1 20
CC CHAIN 21 433
CC CHAIN 21 176
CC CHAIN 155 176
CC CHAIN 178 433
CC DOMAIN 26 62
CC DOMAIN 69 150
CC DOMAIN 151 177
CC DOMAIN 178 433
CC DISULFID 30 38
CC DISULFID 32 50
CC DISULFID 52 61
CC DISULFID 167 298
CC DISULFID 208 224
CC DISULFID 216 287
CC DISULFID 315 384
CC DISULFID 347 363
CC DISULFID 374 402
CC ACT_SITE 223 223
CC ACT_SITE 274 274
CC ACT_SITE 378 378
CC CARBOHYD 324 324
CC SEQUENCE 433 AA; 48595 MW; 816D22DFDDDC8792 CRC64;

Query Match      93.4%; Score 2148; DB 1; Length 433;
Best Local Similarity 92.5%; Pred. No. 1.2e-162;
Matches 389; Conservative 17; Mismatches 10; Indels 4; Gaps 2;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCKPFGGQHCIDKSKTCYEGNGHYRG 60
DB 21 SREL-QVPSDGGCLNGGTCVSNKYFSNIHWCNCKPFGGQHCIDKSKTCYEGNGHYRG 79
QY 61 KASDTMTGRCLPWNSTATVLTQYTHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 120
DB 80 KASDTMTGRSLAWNSATVLTQYTHAHRSDALQLGLGKHNYCRNPDNRPRPCWYVQVGLK 139
QY 121 PLVQECMVHDCADKPKSSPEELKFCQGGKTLRPRFKIIGBEFTTIENQPFWFAAIYRRH 180
DB 140 QRVQECMVHDCADKPKSSPEELKFCQGGKTLRPRFKIIGBEFTTIENQPFWFAAIYRRH 199
QY 181 RGSVTVVCGSLSPCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 240
DB 200 RGSVTVVCGSLSPCWVLSATHCFIDYPKEDYIVVLGSRSLNSNTQGMKEVENLI 259
QY 241 LHKDYSADTLAHNDIALALKIRSKRGCAQPSRTIQTICLPSMYNDPQ---FGTSCBITG 297
```



```

Db 260 LHEDYSADTLAHNDIALKRSREGRCAQPSRTIQTICLPSMYNDPDPFGTSCETIG 319
Qy 298 FGKENSIDLYPEQLKMTWVKLVSHRECOQPHYGVSEVTTKMLCAADPOWKTDSOQDGS 357
Db 320 FGKENSIDLYPEQLKMTWVKLVSHKQCOQPHYGVSEVTTKMLCAADPOWKTDSOQDGS 379
Qy 358 GPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 411
Df 380 GPLVCSIQGHMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGLAL 433

RESULT 3
UROK_PIG ID UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01648; CAA25806.1; -
DR EMBL; X02724; CAA26511.1; -
DR PIR; A00932; UKPG.
DR HSSP; P00749; IKDU.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01186; EGF 2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 287 287
FT CONFLICT 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; EB32FCEP501321EE CRC64;

Query Match 80.8%; Score 1859.5; DB 1; Length 442;
Best Local Similarity 79.3%; Pred. No. 8.3e-140;
Matches 334; Conservative 33; Mismatches 43; Indels 11; Gaps 2;

Qy 1 SNEHQV--PSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIDKSKTCYEGNGHY 58
Db 21 SHEHQBSGASNCGLNGGKCVSYKFSNIQRCSPKKFQGEHCIEDTSQTCFEGNGHSY 80
Qy 59 RGKASTDTMGPPCLPWN SATVLQOITYHAHRS DALQLGLGKHN YCRNPDNRPRPWCYVQVG 118
Db 81 RGKANTNTGGRPCLPWN SATVLLNTYHAHRP DALQLGLGKHN YCRNPDNRPRPWCYVQVG 140
Qy 119 LKPLVQECWHDCA-----DGKFPSSPEELKFCQCGOKTLRPRPKITGGEFTTIE 169
Db 141 LKQLVQECWVNGSGSGSHR PAYDGNKPFSTPEKVEFCQKALRPRFKIVGGKSTTIE 200
Qy 170 QPWFAAIYRRHSGSVTVVCGSLISPCWVTSATHCFIDYPKKEDIYVLGRSLNSNTQ 229
Db 201 QPWFAAIYRRHSGSVTVVCGSLISPCWVTSATHCFINYOQKEDIYVLGRQTLHSSTH 260
Qy 230 GEMKFEVENLILHKDYSDATLAHNDIALKIRSKRGCAQPSRTIQTICLPSMYNDPQF 289
Db 261 GEMKFEVEKLILHEDYSADSLAHNDIALKIRTDKGCAQPSRSIQTICLPPVNGDAHF 320
Qy 290 GTSCEITGFGKENSTDIYLPOLAKMTVVKLISHRECOQPHYGVSEVTTKMLCAADPOWKT 349
Db 321 GASCEIVGFKEDPSDIYLPOLAKMTVVKLVSHRECOQPHYGVSEVTTKMLCAADPOWKT 380
Qy 350 DSCQDSDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGL 409
Db 381 DSCQDSDGGPLVCSLQGRMTLTGIVSWGRGCALKDKPGVYTRVSHFLPWIRSHKTEENGL 440
Qy 410 A 410
Db 441 A 441

RESULT 4
UROK_BOVIN ID UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```


RA Rabbani S.A.;
 RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
 CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
 CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
 CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; X63434; CAA45028.1; -;
 DR EMBL; X63651; CAA46601.1; -;
 DR PIR; S24604; S18932.
 DR HSP; P00749; 1KDU.
 DR MEROPS; S01_231; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR006210; EGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser. Protease_Try.
 DR Pfam; PF00051; Kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Kringle; EGF-like domain; Zymogen; Signal.
 FT SIGNAL 1
 FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
 FT CHAIN 20 177 CHAIN A (BY SIMILARITY).
 FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
 FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
 FT DOMAIN 27 63 EGF-LIKE.
 FT DOMAIN 70 151 KRINGLE.
 FT DOMAIN 152 178 CONNECTING PEPTIDE.
 FT DOMAIN 179 432 SERINE PROTEASE.
 FT DOMAIN 31 39 BY SIMILARITY.
 FT DISULFID 33 51 BY SIMILARITY.
 FT DISULFID 53 62 BY SIMILARITY.
 FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 210 226 BY SIMILARITY.
 FT DISULFID 218 289 BY SIMILARITY.
 FT DISULFID 314 383 BY SIMILARITY.
 FT DISULFID 346 362 BY SIMILARITY.
 FT DISULFID 373 401 BY SIMILARITY.
 FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
 FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
 FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
 FT CONFLICT 16 16 N -> H (IN REF. 2).
 FT CONFLICT 24 24 E -> G (IN REF. 2).
 FT CONFLICT 332 332 D -> N (IN REF. 2).
 SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 73.5%; Score 1690.5; DB 1; Length 432;
 Best Local Similarity 73.2%; Pred. No. 1.9e-126;
 Matches 295; Conservative 47; Mismatches 60; Indels 1; Gaps 1;
 CQ 9 SNCDCLNGTCVSNKYFESNTHWCNCPKFKGQHCHEIDKSKTCYEGNHFYRGKASTDTMG 68
 DB 29 SNOCQNGGVCSYKYFESSRSCPKFKGHCHEIDTSKTCYHNGQSYRGKANTDTKG 88
 QY 69 RPCLPWNASVLOQTYHAHRSALQLGLGHKNYCRNPNRRRPMWYQVGLKPLVQECMV 128
 DB 89 RPCLAWNPAVLQOTYNAHRSALSLGLGHKNYCRNPNQRRPMWYQVGLKQFVQECMV 148
 QY 129 HDCADGKKPSPPEELKFCQCKTLRPRFKIIGGEFTTIENQPMFAAIYRHRGGS-VTY 187
 DB 149 QDCSLSKKPSSTVDQOQFCQKALRPRFKIVGGEFTWENQPMFAAIYLNKNGKSPSP 208
 QY 188 VCGSLSPCWVTSATHCFTDYPKKEDYIVYLGESRLNSNTQGMKEFEVENLILHKDYSA 247
 Ph 209 KCGSLSPCWASATHCFVNPQKKEVYVYVLOQSKNSYNPGEKFEVEQLILHEDFSD 268
 QY 248 DTLAHNDIALLKIRSEKGRCAOPSRITQITCLIPSMYNDPQFGTSCITGFGKENSIDYL 307
 DB 269 ETLAHNDIALLKIRTSTGCAOPSRITQITCLIPREGDAPFGSDCHITFGQESATDYP 328
 QY 308 YPQLQMTVVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTSCQGDSCGSLVCSLQGR 367
 DB 329 YPKDKMSVVVKIISHQCKQPHYGVSEINVKMLCAADPEWKTSCSGDSGGLICNIDGR 388
 QY 368 MTLTGIVSWGRCALDKPQVTRVSHFLPWISHTKEENGLA 410
 DB 389 PTLSGIVSWGSCAEKRNKPGVYTRVSYFLNWIOSHIGEENGLA 431
 RESULT 6
 UROK_MOUSE
 ID UROK_MOUSE STANDARD; PRT; 433 AA.
 AC P06869;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (UPA)
 DE (U-plasminogen activator).
 GN PLAU.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85179474; PubMed=2985383;
 RA Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
 RA Reich E., Kocher H.P., Duvoisin R.M.;
 RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
 RT mouse urokinase-type plasminogen activator";
 RL Eur. J. Biochem. 148:225-232 (1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88163489; PubMed=2831940;
 RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
 RT "The murine urokinase-type plasminogen activator gene";
 RL Biochemistry 26:8270-8279 (1987).
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -


```
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 BY SIMILARITY.
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; BD881048DD666A55 CRC64;

Query Match 43.6%; Score 1004; DB 1; Length 434;
Best Local Similarity 47.4%; Pred. No. 3.7e-72;
Matches 188; Conservative 64; Mismatches 123; Indels 22; Gaps 6;

Oy 11 CDCLEGGFCVSNKFSNHWNCNPKFGCGCEIDKSKTCYEGNGHFYRGKASDTWGRP 70
Db 40 CQCLNGGTCTYRFSQIKRCLCEGYGLHCEIDTNSICYSNGEDYRGWAEDP---G 95
Oy 71 CLPNSATVLO-QTYHAHRSALQGLGKHNCRPNDRPWCVVQGLKPLVQCMVH 129
Db 96 CLYWDHPSVIRGVDYHADLKNALQGLGKHNCRPNDRPWCVVQGLKPLVQCMVH 149
Oy 130 DCADGKPPSPPELKCQGGKTLRPRKIIGFTTEIENPFAIYRHRGSGVTYVC 189
Db 150 -----TPCSTIERCETCGORSFKYKIVGGQAEVETQWIAQINIM-GTDQFLC 202
Oy 190 GGSILPCWVTSATHCFTD---VPKKEDYIVLGRSLNNTQGMKFEVENILHKDY 245
Db 203 GGSILPCWVLTAAHCFNPKPKPNKSVYKVLGKSLINTNDEHQVFWDEIISHPDF 262
Oy 246 SADTLAHNHDIALKIRSKRCQAQPTIQTICLPSMYNDPQFTGCEITFGKENSST 305
Db 263 TDHTGGNDNDIALIRIRATASQCAVESNYRTVCLPEKNLNYDNTWCETIAGYKQNSYD 322
Oy 306 YLYPEQLKMTVVKLISHRECOOPHYGSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQ 365
Db 323 IYFAQLMSATVNLIISQDCNKKYIDSTRVDDNVCAGDPLWETDACKDGGSGGPMVCEHN 382
Oy 366 GRMTLTGVSWGRGKALCKPFGVTVRYSHFLPWIRSH 402
Db 383 GRMTLYGVSWGDCCKAKNKGVTYTRYLNWIDSN 419

RESULT 8
ID _URT2_DESRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSPA
DE alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
```

```
DR SMART; SMO0130; KR; 1.
DR SMART; SMO0020; TRYP_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
FT DOMAIN 40 82 FIBRONECTIN TYPE-1.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE.
FT DOMAIN 225 477 SERINE PROTEASE.
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 214 345 BY SIMILARITY.
FT DISULFID 257 273 BY SIMILARITY.
FT DISULFID 265 334 BY SIMILARITY.
FT DISULFID 359 434 BY SIMILARITY.
FT DISULFID 391 407 BY SIMILARITY.
FT DISULFID 424 452 BY SIMILARITY.
FT CARBOHYD 195 185 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0E5077C CRC64;

Query Match 37.6%; Score 864.5; DB 1; Length 477;
Best Local Similarity 42.9%; Pred. No. 4.4e-61;
Matches 178; Conservative 60; Mismatches 148; Indels 29; Gaps 9;

QY 3 ELHQVP-----SNCDGANGTCVSNKYFSNIHWCNPKPGQGHCEIDKSKTCYEGNGHYF 58
DB 78 QCHTVPVKSCSLRCFNGGTCWQAASFSDF-VCQCPKGYTGKQCEVDTHATCYKQDQGVTY 136
QY 59 RGKASTDTMRCLPNSATVLCQTYHAHRSDALGLGKHNKCNPNRRRPPCYQVVG 118
DB 137 RGTWSTSEGAOCINNSNLLTRTYNGRRSDAITLGLGNHNYCNPNNSKPKWCYVKA 196
QY 119 LKPLVQECMVHDCADGKPPSPPEELFKQCG-QKTLRPRFKIIGGEFTTIENQWFAAIY 177
DB 197 SKFILEFCSPVCS-----KATCGLRKVKKEQLHSTGGLFTDITSHPWQAIF 244
QY 178 RRRHGG- VTYVCGSLISPCWVSIATHCFID- YPKEDYIVYLGSRSLNNTQGMKFE 235
DB 245 AQNRSSGSRFFCGGILITSSCWLVTAACHCFQERYPQHLRVV-LGRTVRVKPGKEEQTFE 303
QY 236 VENLILHKDYSADTLAHHNDIALIKTRSGKRCAPQPSRTIQICLPSMYNDPQFQTSCEI 295
DB 304 VEKCIHVEFDDT--YNNDIALLQLKSGSPCAQESVRAICLPEANILQIPDWTECEL 361
QY 296 TFGKENSTDYLYPEOLKMTVKVLISHRECQPHYVGSVEVTKMLCAADPQMKT----- 349
DB 362 SGYGHKSGSPFYSEQLKEGHRVLPSSRCTSKFLFNKVTNNMLCAGDTRSGEYIPNVH 421
QY 350 DSCQDGGSGPLVCSLQGRMTLGIYSGRGKALCKPGVYTVFVSHFLWIRSHTK 404
DB 422 DACQDGGSGPLVCMNDNHTLGLIISWGVGCGEKDIPGVYTKVTNYLGMIRDNR 476
```

RESULT 9

```
TPA_HUMAN
ID TPA_HUMAN STANDARD; PRT; 562 AA.
AC P00750; Q15103;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) (Alteplase) (Retenplase).
GN PLAT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=83115262; PubMed=6337343;
RA Pennica D., Holmes W.E., Kohr W.J., Harkins R.N., Vehar G.A.,
RA Ward C.A., Bennett W.F., Yelverton E., Seeburg P.H., Heyneker H.L.,
RA Goeddel D.V., Collen D.;
RT "Cloning and expression of human tissue-type plasminogen activator
RT cDNA in E. coli.";
RL Nature 301:214-221(1983).
[2]
RN SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=88262579; PubMed=3133640;
RA Sasaki H., Saito Y., Hayashi M., Otsuka K., Niwa M.;
RT "Nucleotide sequence of the tissue-type plasminogen activator cDNA
RT from human fetal lung cells.";
RL Nucleic Acids Res. 16:5695-5695(1988).
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=88054470; PubMed=2824147;
RA Reddy V.B., Garramone A.J., Sasak H., Wei C.-M., Watkins P., Galli J.,
RA Hsiung N.;
RT "Expression of human uterine tissue-type plasminogen activator in
RT mouse cells using BPV vectors.";
RL DNA 6:461-472(1987).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE=86196143; PubMed=3009482;
RA Friesner Degen S.J., Rajput B., Reich E.;
RT "The human tissue plasminogen activator gene.";
RL J. Biol. Chem. 261:6972-6985(1986).
[5]
RN SEQUENCE FROM N.A.
RX MEDLINE=84298137; PubMed=6089198;
RA Ny T., Elgh F., Lund B.;
RT "The structure of the human tissue-type plasminogen activator gene:
RT correlation of intron and exon structures to functional and
RT structural domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:5355-5359(1984).
[6]
RN SEQUENCE FROM N.A.
RX MEDLINE=86284200; PubMed=3090401;
RA Harris T.J., Patel T., Marston F.A., Little S., Emtage J.S.,
RA Odenakker G., Volckaert G., Rombauts W., Billiau A., Somer P.;
RT "Cloning of cDNA coding for human tissue-type plasminogen activator
RT and its expression in Escherichia coli.";
RL Mol. Biol. Med. 3:279-292(1986).
[7]
RN SEQUENCE FROM N.A. (ISOFORM SHORT).
RC TISSUE=Umbilical vein;
RX MEDLINE=90192129; PubMed=2107528;
RA Siebert P.D., Fong K.;
RT "Variant tissue-type plasminogen activator (PLAT) cDNA obtained from
RT human endothelial cells.";
RL Nucleic Acids Res. 18:1086-1086(1990).
[8]
```


Db 196 GKYSBFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLPWNMSMILIGKYVTAQNP 255
QY 139 -----SPEELK-----FOCG-QKTLRPRFKIGGE 163
Db 256 AQALGLGKHNYCNPDPGADKPCWHLKNNRLTWEYCDVPSCSTGLRQYQPOFRKGG 315
QY 164 FTIENQPFALVRRH-RGGSVTVYCGSLISPCWVISATHCFIDYPKKEDIYVILGRS 222
Db 316 FADIASHPQQAALFAKHRRPGERFLCGGILISCMWLSAAHCFQERPPPHHILTVILGRT 375
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCACQAPSRTIQTICLPS 282
Db 376 YRVVPEEGEEQFEVYKHVEFDDDT--YNDIALILQLKSDSSRCAQESSVVRTCLPP 433
QY 283 MYNDPQFGTSCREITGKKNSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTMLCA 342
Db 434 ADLQLPDWTCELSGVGKHEALSPFYSERLKEARVLYPSRCTSQHLLNRTVTDNMLCA 493
QY 343 AD-----POWKT-DSQCGDSGGPLVCSLOGRMWTLTGIVSWGRCALKDKPGVTVRYSHFL 396
Db 494 GDRSGGPOANLHADACQDSGGPLVCLNDGRMTLVGIISWGLCGGQKQDVPGVTKVTNYL 553
QY 397 PWIRSHTK 404
Db 554 DWIRDNR 561

RESULT 10
URTB_DESRO STANDARD; PRT; 431 AA.
ID P98121;
AC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.69) (DSPA beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland; PubMed=1937019;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M63989; AAA31594.1; -.
DR PIR; JS0599; JS0599.
DR HSP; P98119; IASI.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 431
FT DOMAIN 37 75
FT DOMAIN 82 163
FT DOMAIN 179 431
FT ACT_SITE 226 226
FT ACT_SITE 275 275
FT ACT_SITE 382 382
FT DISULFID 41 52
FT DISULFID 46 63
FT DISULFID 65 74
FT DISULFID 82 163
FT DISULFID 103 145
FT DISULFID 134 158
FT DISULFID 168 299
FT DISULFID 211 227
FT DISULFID 219 288
FT DISULFID 313 388
FT DISULFID 345 361
FT DISULFID 378 406
FT CARBOHYD 139 139
FT CARBOHYD 352 352
SQ SEQUENCE 431 AA; 48221 MW; 699B5E675B162CBF CRC64;

Query Match 37.3%; Score 858.5; DB 1; Length 431;
Best Local Similarity 43.2%; Pred. No. 1.2e-60;
Matches 175; Conservative 59; Mismatches 146; Indels 25; Gaps 8;

QY 9 SNCCDLNGGTCVSNKYFSNIHNCNPKFGGQHCIDKSKTCYEGNGHYRGKASTDTMG 69
Db 42 SEILCPNGGTCWQAASFSDF-VCCQPKGYTKQCEVDTHATCYKQDQGVYRGWTSTSEG 100
QY 69 RECLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDRNRPPWCYVQGLKPLVQECMV 128
Db 101 AQCIWN SNLLRTYNGERSDAITLGLGNHNYCRNPDDNSKPCWCVIKASFILFCV 160
QY 129 HPCADKGPSPPELKFQCG-QKTLRPRFKIIGGEFTTIENQPFALVRRHGRGS-VT 186
Db 161 PVCS-----KATCGLRKYKEPOLHSTGGLFTDITSHPQQAALFAQNRSSGER 208
QY 187 YVCGSLISPCWVISATHCFID-YPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDY 245
Db 187 YVCGSLISPCWVISATHCFID-YPKKEDIYVILGRSLNSNTQGMKFEVENLILHKDY 245


```
Db 209 FLOGGILISSCWLTAAHCFQERYPPQHLRV-LGRTYRVKPKBQTFEVEKCIHEEF 267
Qy 246 SADTLAHNDIALKIRSKGRCAQPSRTIQTICLPMSYNDPQFTSCETGFKENSTD 305
Db 268 DDDT--YNNDIALLQLKSPQCAQESDSVRAICLPEANLQLPDWTCELSGVGKHKSS 325
Qy 306 YLYPEQLKMTVVKLISHRECCQPHYGVSEVTTKMLCAADPQWKT-----DSQCQDSGGP 359
Db 326 PFYSEQLKEGHVRLYPSRSTKFLFNKVTNNMLCAGDTRSGEITYPNVHDAQCQDSGGP 385
Qy 360 LVCSLQGRMTLTGIVSGRGKALKDQGVTRYVSHFLPWIRSHTK 404
Db 386 LVCNDNHTLLGIISGVGCKEDIFGVTKVNYLGLWIRDNR 430

RESULT 11
TPA RAT
ID TPA RAT STANDARD; PRT; 559 AA.
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, last sequence update)
DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Tissue-type plasminogen activator precursor (RC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89170114; PubMed=3148445;
RA My T.; Leonardson G.; Hsueh A.J.W.;
RT "Cloning and characterization of a cDNA for rat tissue-type
RT plasminogen activator."
RL DNA 7:671-677(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130448; PubMed=2105315;
RA Feng P.; Ohlsson M.; Ny T.;
RT "The structure of the TAT-1 rat tissue-type plasminogen activator
RT gene. Species-specific sequence divergences in the promoter predict
RT differences in regulation of gene expression."
RL J. Biol. Chem. 265:2022-2027(1990).
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL; M23697; AAA41812.1; -.
```

```
DR EMBL; M31197; AAA42261.1; -.
DR EMBL; M31185; AAA42261.1; JOINED.
DR EMBL; M31186; AAA42261.1; JOINED.
DR EMBL; M31187; AAA42261.1; JOINED.
DR EMBL; M31188; AAA42261.1; JOINED.
DR EMBL; M31189; AAA42261.1; JOINED.
DR EMBL; M31190; AAA42261.1; JOINED.
DR EMBL; M31191; AAA42261.1; JOINED.
DR EMBL; M31192; AAA42261.1; JOINED.
DR EMBL; M31193; AAA42261.1; JOINED.
DR EMBL; M31194; AAA42261.1; JOINED.
DR EMBL; M31195; AAA42261.1; JOINED.
DR EMBL; M31196; AAA42261.1; JOINED.
DR EMBL; A19618; CAA01482.1; -.
DR PIR; A35029; A35029.
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SMC0020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SRR; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT CHAIN 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 36 78 CHAIN.
FT DOMAIN 36 78 FIBRONECTIN TYPE-I.
FT DOMAIN 79 117 EGF-LIKE.
FT DOMAIN 124 205 KRINGLE 1.
FT DOMAIN 213 294 KRINGLE 2.
FT DOMAIN 309 559 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 65 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
```


FT	SIGNAL	1	36	POTENTIAL.
FT	CHAIN	37	477	SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.
FT	DOMAIN	40	82	FIBRONECTIN TYPE-I.
FT	DOMAIN	83	121	EGF-LIKE.
FT	DOMAIN	128	209	KRINGLE.
FT	DOMAIN	225	477	SERINE PROTEASE.
FT	ACT_SITE	272	272	CHARGE RELAY SYSTEM.
FT	ACT_SITE	321	321	CHARGE RELAY SYSTEM.
FT	ACT_SITE	428	428	CHARGE RELAY SYSTEM.
FT	DISULFID	42	72	BY SIMILARITY.
FT	DISULFID	70	79	BY SIMILARITY.
FT	DISULFID	87	98	BY SIMILARITY.
FT	DISULFID	92	109	BY SIMILARITY.
FT	DISULFID	111	120	BY SIMILARITY.
FT	DISULFID	128	209	BY SIMILARITY.
FT	DISULFID	149	191	BY SIMILARITY.
FT	DISULFID	180	204	BY SIMILARITY.
FT	DISULFID	214	345	BY SIMILARITY.
FT	DISULFID	257	273	BY SIMILARITY.
FT	DISULFID	265	334	BY SIMILARITY.
FT	DISULFID	359	434	BY SIMILARITY.
FT	DISULFID	391	407	BY SIMILARITY.
FT	DISULFID	424	452	BY SIMILARITY.
FT	CARBOHYD	153	153	N-LINKED (GLCNAC. .).
FT	CARBOHYD	398	398	/FTId=CAR_000027.
FT	CARBOHYD	398	398	N-LINKED (GLCNAC. .).
FT	CARBOHYD	398	398	/FTId=CAR_000028.
FT	TURN	214	215	
FT	STRAND	223	224	
FT	TURN	226	227	
FT	STRAND	230	231	
FT	HELIX	234	236	
FT	TURN	238	239	
FT	STRAND	240	245	
FT	STRAND	254	263	
FT	TURN	264	265	
FT	STRAND	266	269	
FT	HELIX	271	273	
FT	TURN	280	282	
FT	STRAND	284	287	
FT	TURN	297	298	
FT	STRAND	300	309	
FT	TURN	311	312	
FT	TURN	315	317	
FT	TURN	319	320	
FT	STRAND	323	328	
FT	STRAND	338	338	
FT	TURN	339	341	
FT	STRAND	341	341	
FT	STRAND	345	345	
FT	TURN	349	350	
FT	TURN	355	356	
FT	STRAND	358	363	
FT	STRAND	366	366	
FT	STRAND	374	374	
FT	STRAND	379	385	
FT	STRAND	388	390	
FT	HELIX	393	398	
FT	TURN	403	404	
FT	STRAND	405	409	
FT	TURN	425	426	
FT	TURN	428	429	
FT	STRAND	431	436	
FT	TURN	437	438	
FT	STRAND	439	448	
FT	TURN	455	456	
FT	STRAND	459	463	
FT	STRAND	464	467	
FT	HELIX	468	474	
SEQ	SEQUENCE	477 AA;	53616 MW; AA06PD1739C10B5E CRC64;	

Query Match 36.4%; Score 836.5; DB 1; Length 477;
Best Local Similarity 42.2%; Pred. No. 7.2e-59;

Matches	175;	Conservative	56;	Mismatches	147;	Indels	37;	Gaps	10;
QY	5	HQVPSN-CD---	CLNGGTCVSNKYFSNIHWCNCPKKGQCHCEIDKSKTYEGNGHFYRG	60					
DB	80	HTVPNSCSEPRCFNGGTCWQAVYFSDP	-VQCQAGYTGKCEVDTRATCEGQGVYRG	138					
QY	61	KASDTMGRPCLPNWSATVLQOTYHAHRSDALQGLGKHNYCRPNRRRRCWYVOVGLK	120						
DB	139	TWSTAESRVEICINWNSLLTRTYNGRMPDAFNGLGNHNYCRNPGAPKWCYVIKAGK	198						
QY	121	PLVQECMVHDCADGKSSPEELKFCQG	-QKTLRPREFKIIGGBPTTIENOPWFAAIYR	179					
DB	199	FTSSCSVPVCS-----	KATCGLRKYKEPQLHSTGGUFTDITSHPWQAIAFAQ	246					
QY	180	HRGGS-VTVYCGSLISPCWVISATHCFIDYPPKEDYI	-----VYLGSRSLNSNTQGEEMK	233					
DB	247	NRRSSGERFLCGGILISSCWVLTAAHCF	-----QESYLPDQLKVGLGTYRVKGEBOQT	301					
QY	234	FEVENLILHKDYSADTLAHNDIALKIRSKGRCQAOPSRTIOTICLPSMYNDPOFGTSC	293						
DB	302	FKVKYIVHKEFDDDT--YNNDIALQLKSDSPQCAQESDSVRAICLPEANLQLPDWTEC	359						
QY	294	BITGFGKENSTDYLYPEOLKMTVVKLISHRECQQPHYYGSEVTKMLCAADPQWKT	--- 349						
DB	360	ELSGYGHKSSSPYSEQLKEGHVLYPSSRCAPKFLFNKVTNNMLCAGDTRSGEIVPN	419						
QY	350	--DSCQDSDGGPLVCSLQGRMTLTGIVSWRGCAKDKPGVYTVRVSHPFPIRSH	402						
DB	420	VHDACQDSDGGPLVCMNDNHTLGIISWGVCGEKDVPGVYTKVTNYLGIWRDN	474						
RESULT 13									
ID TPA_MOUSE STANDARD; PRT; 559 AA.									
AC	P11214; O91VP2;								
DT	01-JUL-1989 (Rel. 11, Created)								
DT	15-SEP-2003 (Rel. 42, Last sequence update)								
DT	15-SEP-2003 (Rel. 42, Last annotation update)								
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)								
DE	(t-PA) (t-plasminogen activator).								
GN	PLAT.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxID=10090;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RP	MEDLINE=88087303; PubMed=2826484;								
RA	Rickles R.J., Darrow A.L., Strickland S.;								
RT	"Molecular cloning of complementary DNA to mouse tissue plasminogen								
RT	activator mRNA and its expression during F9 teratocarcinoma cell								
RT	differentiation.";								
RL	J. Biol. Chem. 263:1563-1569(1988).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Mammary gland;								
RA	MEDLINE=22388257; PubMed=12477932;								
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,								
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,								
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,								
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,								
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,								
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,								
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,								
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,								
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,								
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,								
RA	Faney J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,								
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,								
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,								
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,								
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,								

RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; J03520; AAA0470.1; --
CC ENBL; BC011256; AAH1256.1; --
CC PLK; A29941; A29941.
CC HSP; P00750; IA5H.
CC MEROPS; S01.232; --
CC MGD; MGI:97610; Plat.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fnl; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FNI; 1.
CC SMART; SM00130; KR; 2.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF 1; 1.
CC PROSITE; PS01186; EGF 2; 1.
CC PROSITE; PS01253; FIBRONECTIN 1; 1.
CC PROSITE; PS00021; KRINGLE 1; 2.
CC PROSITE; PS00070; KRINGLE 2; 2.
CC PROSITE; PS0240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC plasminogen activation: Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17
FT PROPEP 18 29 PROBABLE.
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 CHAIN.
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
FT CHAIN 309 559 CHAIN.
FT DOMAIN 36 78 FIBRONECTIN TYPE-I.

FT	DOMAIN	79	117	EGF-LIKE.
FT	DOMAIN	124	205	KRINGLE 1.
FT	DOMAIN	213	294	KRINGLE 2.
FT	DOMAIN	309	559	SERINE PROTEASE.
FT	ACT_SITE	355	355	CHARGE RELAY SYSTEM.
FT	ACT_SITE	404	404	CHARGE RELAY SYSTEM.
FT	ACT_SITE	510	510	CHARGE RELAY SYSTEM.
FT	DISULFID	38	68	BY SIMILARITY.
FT	DISULFID	66	75	BY SIMILARITY.
FT	DISULFID	83	94	BY SIMILARITY.
FT	DISULFID	88	105	BY SIMILARITY.
FT	DISULFID	107	116	BY SIMILARITY.
FT	DISULFID	124	205	BY SIMILARITY.
FT	DISULFID	145	187	BY SIMILARITY.
FT	DISULFID	176	200	BY SIMILARITY.
FT	DISULFID	213	294	BY SIMILARITY.
FT	DISULFID	234	276	BY SIMILARITY.
FT	DISULFID	265	289	BY SIMILARITY.
FT	DISULFID	297	428	INTERCHAIN (BY SIMILARITY).
FT	DISULFID	340	356	BY SIMILARITY.
FT	DISULFID	348	417	BY SIMILARITY.
FT	DISULFID	442	516	BY SIMILARITY.
FT	DISULFID	474	490	BY SIMILARITY.
FT	DISULFID	506	534	BY SIMILARITY.
FT	CARBOHYD	149	149	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	260	260	G -> A (IN REF. 1).
FT	CONFLICT	325	325	P -> A (IN REF. 1).
SQ	SEQUENCE	559 AA;	63122 MW;	8CCBE2BDB94514D9 CRC64;

Query Match 35.9%; Score 826.5; DB 1; Length 559;
Best Local Similarity 35.7%; Pred. No. 5.3e-59;
Matches 176; Conservative 65; Mismatches 155; Indels 97; Gaps 10;

QY	3	ELHOVP----	SNCDCLNGGTCVSNKYFSNTHWNCNPKFGGCHCEIDKSKTCYEGNGHFY	58
Db	74	QCHSVPRVSCSEPRCNGGTCQALFSDP	VCQPDGFVGKCRDIATRCFEEQITY	132
QY	59	RGKASTDTMGRPCLPWNSATVLOOTYAHASDALQLGLGKHNYCRNPNRRRPMWYQVG	118	
Db	133	RGTWSTAESGAECINWNSSVLSLKPYNARRPNAIKLGLGNHNYCRNPNDRDLKPCWYFKA	192	
QY	119	LKPLVQECMVHDCADGKPS-----	138	
Db	193	GKYTEFCSTPACPKGKSEDCYVGKGYVYRGTHSLTTSQASCLPWNSIVLMGKSYTAWRT	252	
QY	139	-----SPEELK-----	162	
Db	253	NSQALGLGRHNYCRNPDGDARPMCHWKORKLWECYCDMSPCSTCGURQYKRPQFRKGG	312	
QY	163	EFTTIENQPMFAAIY-RRHRGGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGR	221	
Db	313	LYTDITSHWPQAPIFVKNKRSPGERFLCGGLISSCWLSAAHCFLEFPFNHLKVVLR	372	
QY	222	SRLNSNTQGMKEFEVENLILHKDYSADTLAHNDIALKLRSEKGRCAQPSRTIOTICLP	281	
Db	373	TYRVFGEBEQTFEIEKYIVHEFFDDT--YNDNIALQLRSQSKCAQESSVGTACLP	430	
QY	282	SMYNDPQF-----GTSCEITGFGKENSTDYLYPBOLKMTVVVKLISHRECCQPHYVGSVTT	337	
Db	431	-----DENLQLPDWTECELSGYGKHEASSFFSDRLKEAHVRLYPSRSCRTSQHLFNKVTVN	486	
QY	338	KMLCAADP-----QWKTDSCQGDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRV	392	
Db	487	NMLCAGDTRSGGNQDLHDACQGDGGPLVCMINKMTLTGIIISWGLGCGQKDPGVYTKV	546	
QY	393	SHPLPWIRSHTKK	405	
Db	547	TNYLDWIHDNMKQ	559	

RESULT 14
TPA_BOVIN

ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAI.
OS Bos taurus (Bovine).
OC Bakayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
and tPA";
RL Int. Dairy J. 5:605-617 (1995).
CC -1- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZIMOGEN PLASMINOGEN
TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
MANY OTHER PHYSIOLOGICAL EVENTS.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -1- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -1- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
ARG-314 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
CC -1- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 2 kringle domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X85800; CAA53795.1; -.
CC HSSP; P00750; IRTF.
CC MEROPS; S01.232; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000083; Fibrinctnl.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00051; kringle; 2.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 2.
CC SMART; SM00181; EGF; 1.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00130; KR; 2.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF-2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 2.

DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT PROPEP 22 33 BY SIMILARITY.
FT CHAIN 34 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 34 314 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
CHAIN.
FT CHAIN 315 566 TISSUE-TYPE PLASMINOGEN ACTIVATOR B
CHAIN.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE 1.
FT DOMAIN 219 300 KRINGLE 2.
FT DOMAIN 315 566 SERINE PROTEASE.
FT ACT_SITE 361 361 CHARGE RELAY SYSTEM.
FT ACT_SITE 410 410 CHARGE RELAY SYSTEM.
FT ACT_SITE 517 517 CHARGE RELAY SYSTEM.
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 219 300 BY SIMILARITY.
FT DISULFID 240 282 BY SIMILARITY.
FT DISULFID 271 295 BY SIMILARITY.
FT DISULFID 303 434 INTERCHAIN (BY SIMILARITY).
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 354 423 BY SIMILARITY.
FT DISULFID 448 523 BY SIMILARITY.
FT DISULFID 480 496 BY SIMILARITY.
FT DISULFID 513 541 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 487 487 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BE4E32276C3 CRC64;
Query Match 35.5%; Score 818; DB 1; Length 566;
Best Local Similarity 36.5%; Pred. No. 2.5e-57;
Matches 179; Conservative 62; Mismatches 158; Indels 92; Gaps 11;
QY 3 ELHQP- SNCD- ---CLNGTGVSNKYPFNIHWCNPKKFGQCEIDKSKTCYEGNGHY 58
DB 78 QCHSVPRVSCSEPCWFCNGGTCRQALYSSDF- VQCPEGFMGKLCEDATATCYKDGVA 136
QY 59 RGKASTDTMGRPCLPWN SATVLQOYTHAHRSDALQLGLGHKHNCRPNRRRWCYVQVG 118
DB 137 RGTWSTAESGAECANWNSGLAMKPYSGRRPNAIRLGLGNHNYCRNPDQDSKEWYVKA 196
QY 119 LKPLVQBCMVHDC- ---DGK- --- 135
DB 197 GKVISFCSTPACAKVAEEDGDCYTGNGLAYRTRSHTSKSGASCLPWNVSFLT SKIYTAW 256
QY 136 KPSSP- ---PEELKFQ- ---CG-QKTLRPRFKII 160
DB 257 KSNAPALGLGKHNCRNPDGAQPCWCHVWKDRLTWECYCDVPCVTCGLRQYKFPQRIK 316
QY 161 GGEFTTIENQWFAAIY-RRHGGSVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYL 219
DB 317 GGLFADITSHPMQAAIFVKNRRSPGERFLCGGILISSCWVLSAAHCFQERYPPHLLKVFL 376
QY 220 GRSLSNSNTQGMKFEVENILILHKDYSADTLAHNDIALLLKIRSKGRCAQSRITQIC 279
DB 377 GRTYRLVPGEEQTFEVEKIIHKFDDDT- -YNDIALLLHLKSLTCAESASVETIC 434
QY 280 LPSMYNDPQGTSCETGFGKENSTDYLYPEOLKMTVVKLISHRECQPHYTGSEVTTKM 339
DB 435 LPDASLQLPDWTECELSGVGKHSSPFPFSERLKEAHVRLYPSRCTS QHLFNRTVTNNM 494

```

QY 340 LCAADPOM-----KTDSQGGSGGGLVCSLQGRMTLTGIYSGRGCGALKDKPGVYTRVS 393
DB 495 LCAGTRGGGHTNLHDACQGGSGGLVCMKDNHMLVGLISWGLGCGKQDVGVTYTKVT 554
QY 394 HPLWIRSHK 404
DB 555 NYLDWIRDNR 565

RESULT 15
ID URTG DESRO STANDARD; PRT; 394 AA.
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA
DE gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Mitt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M63990; AAA31595.1; -.
CC PIR: JS0600; JS0600.
CC HSP: P98119; IAS1.
CC MEROPS: S01.239; -.
CC InterPro: IPR001314; Chymotrypsin.
CC InterPro: IPR000001; Kringle.
CC InterPro: IPR001254; Ser_protease_Try.
CC Pfam: PF00051; Kringle; 1.
CC Pfam: PF00089; trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC PRINTS: PR00018; KRINGLE.
CC ProDom: PD000395; Kringle; 1.
CC SMART: SM00130; KR; 1 Spc; 1.
CC SMART: SM00020; Tryp Spc; 1.
CC PROSITE: PS00021; KRINGLE_1; 1.

```

```

DR PROSITE; PSS0070; KRINGLE_2; 1.
DR PROSITE; PSS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 45 126 KRINGLE.
FT DOMAIN 142 394 SERINE PROTEASE.
FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match 32.7%; Score 752; DB 1; Length 394;
Best Local Similarity 42.0%; Pred. No. 2.8e-52;
Matches 155; Conservative 56; Mismatches 134; Indels 24; Gaps 7;

QY 45 DKSITCYEGNGHYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRN 104
DB 40 DPHATCYKQGVYRGTWSTSESQAQCINMNSNLLIRTYNGRMPEAVKLGNGHNYCRN 99
QY 105 PDNRPRWCYVQVGLKPLVQCMVHDCADGKPSPPPEELKFCQG-KQTLRPRFKIIIGE 163
DB 100 PDGASKPCWCYVIKARKTSESCTSPVCS-----KATCLRYKEPQLHSTGGL 147
QY 164 PTTTENQPFWAAIYRRHRGGS-VTVVCGGSLISPCWISATHCFID-YPKEDYIVYLGR 221
DB 148 FDTITSHPWQAAIFAQNRSSGERFLOGGILISSCWLTAAHCFQERYPPQHLRVV-LGR 206
QY 222 SRLNSNTGEMKFEVENILHDKYSADTLAHNDIALKIRSKGRCQAPRTQTICLP 281
DB 207 TYRVKPKGEQTFEVEKIVHEEFDDDT--YNNDIALLQKSGSPQCAQESDRAICLP 264
QY 282 SWYNDPQGTSCETGFGKENSTDYLPQOLKMTWKLIHRECOQPHYVGSEVTTKMLC 341
DB 265 EANLQLPDWTECELSGYGKHKSSPFYSEQLKEGHVRLYPSRCTSKFLFKNTVTNNMLC 324
QY 342 AADPOWKT-----DSQGSQGGPLVCSLQGRMTLTGIYSGRGCGALKDKPGVYTRVSHF 395
DB 325 AGDTRSGEIVPNVHDACQGGSGGLVCMNDNMHTLLGIISWGLGCGKDIPIGVYTKVTNY 384
QY 396 LPWIRSHK 404
DB 385 LGWIRDNR 393

RESULT 16
HGFA HUMAN
ID HGFA HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7693665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028(1993).
[2]
RN SEQUENCE OF 40-655 FROM N.A.
RP
RA Zhao S., Odell C.;
RD Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -!- TISSUE SPECIFICITY: LIVER.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D14012; BAA03113.1; -;
CC EMBL; Z69923; CAA93803.1; -;
CC PIR; A46688; A46688.
CC HSP; P00763; IDPO.
CC MEROPS; S01.228; -;
CC Genew; HGNC:4894; HGFAC.
CC MIM; 604552; -;
CC GO; GO:0005576; C:extracellular; TAS.
CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR000083; Fibrinctn1.
CC InterPro; IPR000562; FN Type II.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_protease_Try.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00039; fn1; 1.
CC Pfam; PF00040; fn2; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PF00722; CHYMOTRYPSIN.
CC PRINTS; PF00013; FNYPEII.
CC PRINTS; PF00018; KRINGLE.
CC ProDom; PD000995; FN Type II; 1.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00058; FN1; 1.
CC SMART; SM00059; FN2; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp SPC; 1.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS01253; FIBRONECTIN_1; 1.
CC PROSITE; PS00023; FIBRONECTIN_2; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
FT EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 30
FT PROPEP 31 372 CLEAVED IN ACTIVE FORM.
FT CHAIN 373 407 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT CHAIN 408 655 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT CHAIN 408 655 CHAIN.
FT DOMAIN 108 148 FIBRONECTIN TYPE-II.
FT DOMAIN 160 198 EGF-LIKE 1.
FT DOMAIN 200 240 FIBRONECTIN TYPE-I.
FT DOMAIN 241 279 EGF-LIKE 2.
FT DOMAIN 286 367 KRINGLE.
FT DOMAIN 408 655 SERINE PROTEASE.
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 598 598 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 108 133 BY SIMILARITY.
FT DISULFID 122 148 BY SIMILARITY.
FT DISULFID 164 175 BY SIMILARITY.
FT DISULFID 169 186 BY SIMILARITY.
FT DISULFID 188 197 BY SIMILARITY.
FT DISULFID 202 230 BY SIMILARITY.
FT DISULFID 228 237 BY SIMILARITY.
FT DISULFID 245 256 BY SIMILARITY.
FT DISULFID 250 267 BY SIMILARITY.
FT DISULFID 269 278 BY SIMILARITY.
FT DISULFID 286 367 BY SIMILARITY.
FT DISULFID 307 349 BY SIMILARITY.
FT DISULFID 338 362 BY SIMILARITY.
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
FT DISULFID 432 448 BY SIMILARITY.
FT DISULFID 440 510 BY SIMILARITY.
FT DISULFID 535 604 BY SIMILARITY.
FT DISULFID 567 583 BY SIMILARITY.
FT DISULFID 594 622 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 644 644 R -> Q (IN REF. 2).
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B862ED7 CRC64;

Query Match 32.0%; Score 735.5; DB 1; Length 655;
Best Local Similarity 37.6%; Pred. No. 1e-50;
Matches 160; Conservative 53; Mismatches 166; Indels 47; Gaps 9;

QY 5 HQVPSNCDCLNGGTCVSNKYFSNTHW-----CNCPEKFGQHCETDKSKTCYEGNGH 56
Db 242 HTACLSPPCLNGGTC-----HLIVATGTTVCACPPGAGRLCNTEPDERCFLNGT 292
QY 57 FYRGKASTDTMGREPLPWN SATVLQQTYYAHRS DALQLGLGKHNVCNPNRRRPPWCYVQ 116
Db 293 GYRGVASTSASGLSCLAWNSDLLYQLHVDVSGAALLGLGPHAYCRNPNDRPWCYV 352
QY 117 VGLAPLVQECMVHDC-----ADGKKPSSPEELKFCQGX-----TLRPRFK 158
Db 353 KDSALSWEYCRLEACE SLTRVQLSPDLLATLPPEASPGRQ-----ACGRHKRKTFLRPR-- 407
QY 159 IIGGEFTTIENQPFPAALYRHRGSGVYVCGSLISPCWVISATHCFIDYPKEDIVY 218
Db 408 IIGSSSLPGSHPWIAAIY---IGDS---FCAGSLVHTCVWVSAACHFSPPRDSVSV 461
QY 219 IGRSLNSNTQGENKMFVENLILHKDYSADT LAHNDIALKTRSKGRCQAQPSRTIQT 278
Db 462 LGQHFFNRTTDTVTQTGIEKVIPTLYSVFNPSDH-DLVLRLLKKGDRCATRSQFQPI 520
QY 279 CLPSMYNDPQFGTSCEITGFGKENSTDYLPQOLKMTVVKLISHRECQPHYYSSEVTTK 338


```
Db 521 CUPEPGSTPAGHKCQAGHLDENVGVSSLRALVPLVADHKCSSPEVYGADISPN 580
Qy 339 MLCAADPQWKTDSCQSDGSLVCSLQGRMTLGIIVSWGRGCALDKPKGVYTVRVSHFLPW 398
Db 581 MLCAGYFDCKSDACQSDGSLVCSLQGRMTLGIIVSWGRGCALDKPKGVYTVRVSHFLPW 640
Qy 399 IRSHTK 404
Db 641 INDIR 646

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC Q9R098; Q9JKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
activator) (HGFA).
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RT Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for
metanephric kidney morphogenesis in vitro.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; AF099017; AAF02489.1; -.
DR EMBL; AF224724; AAF34712.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.228; -.
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
```

```
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00049; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPsin.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SRR; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
CHAIN.
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
CHAIN.
FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 284 BY SIMILARITY.
FT DISULFID 286 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
FT DISULFID 335 359 INTERCHAIN (BY SIMILARITY).
FT DISULFID 392 519 BY SIMILARITY.
FT DISULFID 430 446 BY SIMILARITY.
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;
```

Query Match 31.6%; Score 726; DB 1; Length 653;
Best Local Similarity 37.3%; Pred. No. 5.6e-50;
Matches 158; Conservative 57; Mismatches 167; Indels 42; Gaps 9;

Qy

5 HQVPSNCDLNGGTCVSNKYFSNIHW-----CNCPKKFGGQHCEIDKSKTCYEGNGH 56


```
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DCB946FB9ED59 CRC64;

Query Match 31.2%; Score 718.5; DB 1; Length 603;
Best Local Similarity 36.8%; Pred. No. 2e-49;
Matches 161; Conservative 64; Mismatches 149; Indels 63; Gaps 13;

Qy 13 CLNGCTVSNKYFNIHWCNCPKFGGQCHCIBDKSKTCYEGNGHFYRGKASTDTMGRECL 72
Db 182 CLNGCRCL-VEGHLLCDPCMGYTPGCDLDTTASCYEGRGVSYRGMAITTVSGAKCQ 238

Qy 73 PWSATVLQOYTHAHRSD-ALQGLGKHNYCRNPNRRRRPWCYVOVGLKPLVQSRVHDC 131
Db 239 RWAS-EATYRNNTAQALRRGLGHHTFCRPNDRTPWCFYVMGNRLSWBYCDLAQC 294

Qy 132 ADGKPPSPPELKFQ-----CGOKTLRPR 156
Db 295 QYPPQPTATPHD-RPEHKLPSRLSILQPTQTONALANELPETSSLLCGOR-LRKR 352

Qy 157 F---KIIGBEFTTIENQFPAALYRRHRGGSVTVCGSLISPCWVISATHCFIDYPKK 212
Db 353 LSSLRIYVGLVALPGAHPYIAALY---WGS--NFCGSLIAPCWVLTAAHCLQNRFP 406

Qy 213 EDYIVYLGSRSLNSNTQEMKFEVENLILHKDYSADTLAHNDIALLKI-RSKEGRCAOP 271
Db 407 BELKVLQCDRHNSCEHCQTLAVHSYRLHAFSPS--SYLNDLALLRLQKSDAGSCAOL 464

Qy 272 SRTIQTICLPWYNDPQFG--TSCEITGFGKENSTDYLYPQLKMTVVKLISHRECOQPH 329
Db 465 SPYVTVCLPSGAPPSSESTTCCEVAGCQHGFEGABEYSSFLQEAQVPLISSERCSSPE 524

Qy 330 YGSEVVTMLCAADPQWKTSCQDGGPLVC---SLQGRMTLTGIVSWGRGCAKDKP 386
Db 525 VHGDAFLSGMLCAGLEGTTDACQDGGPLVCDEAAEHLRLIRGIVSWGSGGDRNKP 584

Qy 387 GYTVRVSHFLPWRSH 403
Db 585 GYITDVASILTQIKHT 601

RESULT 19
ID FA12 HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
OS F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGilivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene. Intron/exon gene organization and analysis of the 5'-flanking region.";
RT J. Biol. Chem. 262:13662-13673 (1987).
[2]
RN SEQUENCE FROM N.A., AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldane S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
```

```
RN SEQUENCE OF 4-615 FROM N.A.
RP MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RN Nucleic Acids Res. 14:3146-3146(1986).
[4]
RN SEQUENCE OF 14-615 FROM N.A.
RP MEDLINE=86033930; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGilivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA. Prediction of the primary structure of factor XII and the tertiary structure of beta-factor Xlla.";
RL J. Biol. Chem. 260:13666-13676(1985).
[5]
RN SEQUENCE OF 146-615 FROM N.A.
RP MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman factor).";
RL Biochemistry 25:1525-1528(1986).
[6]
RN SEQUENCE OF 20-379.
RP MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor Xlla (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
[7]
RN SEQUENCE OF 354-362 AND 373-615.
RP MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor Xlla.";
RL J. Biol. Chem. 258:10924-10933(1983).
[8]
RN SEQUENCE OF 561-615 FROM N.A.
RP TISSUE=Blood;
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G->A) in the factor XII gene causes a truncated transcript in cross-reacting material negative patients.";
R Hum. Mol. Genet. 4:1235-1237(1995).
[9]
RN CARBOHYDRATE-LINKAGE SITE THR-109.
RP MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
[10]
RN VARIANT WASHINGTON D.C. SER-590.
RP MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive factor Xlla results from Cys-571-->Ser substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
[11]
RN VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;
RA Hovig J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,
RA Laemmle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by the amino acid substitution Arg-353-->Pro leading to loss of a kallikrein cleavage site.";
RL Blood 84:1173-1181(1994).
[12]
RN VARIANT TENRI CYS-53.
RP MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
```

"Factor XII Tenri, a novel cross-reacting material negative factor XII deficiency, occurs through a proteasome-mediated degradation.";
 RL Blood 93:4300-4308(1999).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-lle bonds in factor VII to form factor VIIa and factor XI to form factor XIa.
 CC -!- PM: O- AND N-GLYCOSYLATED.
 CC -!- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR XIIA ACTIVATES FACTOR XI TO FACTOR XIa.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M31315; AAA70225.1; -;
 DR EMBL; AF538691; AAM97932.1; -;
 DR EMBL; M11723; AAA51986.1; -;
 DR EMBL; M17466; AAB59490.1; -;
 DR EMBL; M17464; AAB59490.1; JOINED.
 DR EMBL; M17465; AAB59490.1; JOINED.
 DR EMBL; M13147; AAA70224.1; -;
 DR EMBL; U71274; AAB51203.1; -;
 DR PIR; A29411; KFHU12.
 DR HSSP; P00763; 1DPO.
 DR MEROPS; S01.211; -;
 DR GENE; HGNC:3530; F12.
 DR MIN; 234000; -;
 DR GO; GO:0003805; F:blood coagulation factor XI activity; TAS.
 DR GO; GO:0003806; F:blood coagulation factor XII activity; TAS.
 DR GO; GO:0007596; P:blood coagulation; TAS.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR Pfam; PF0001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fni; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
 KW Polymorphism; Disease mutation.
 FT SIGNAL 1 19
 FT CHAIN 20 372 ALPHA-FACTOR XIIA HEAVY CHAIN.
 FT CHAIN 373 615 ALPHA-FACTOR XIIA LIGHT CHAIN.
 FT CHAIN 354 362 BETA-FACTOR XIIA PART 1.
 FT CHAIN 373 615 BETA-FACTOR XIIA PART 2.
 FT DOMAIN 47 88 FIBRONECTIN TYPE-II.
 FT DOMAIN 94 131 EGF-LIKE 1.
 FT DOMAIN 133 173 FIBRONECTIN TYPE-I.
 FT DOMAIN 174 210 EGF-LIKE 2.
 FT DOMAIN 217 295 KRINGLE.
 FT DOMAIN 296 349 PRO-RICH.
 FT DOMAIN 373 615 SERINE PROTEASE.
 FT CARBOHYD 109 109 O-LINKED (FUC).
 FT CARBOHYD 249 249 O-LINKED (GLCNAC...).
 FT CARBOHYD 299 299 O-LINKED (POTENTIAL).
 FT CARBOHYD 305 305 O-LINKED (POTENTIAL).
 FT CARBOHYD 308 308 O-LINKED (POTENTIAL).
 FT CARBOHYD 328 328 O-LINKED (POTENTIAL).
 FT CARBOHYD 329 329 O-LINKED (POTENTIAL).
 FT CARBOHYD 337 337 O-LINKED (POTENTIAL).
 FT ACT_SITE 412 412 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 461 461 CHARGE RELAY SYSTEM (BY SIMILARITY).
 Query Match 30.1%; Score 692; DB 1; Length 615;
 Best Local Similarity 34.3%; Pred. No. 2.6e-47;
 Matches 153; Conservative 60; Mismatches 163; Indels 70; Gaps 9;
 QY 13 CLNGTGVSNKYFSNIHWCNCPKFGGHOCHIDSKTCYEGNGHFGKASTDMGRCL 72
 DB 183 CLHGRCLC---VEGHLCHCPGVGYTGFCDVTKASYDGRGLSYRGLARTTISGAPCQ 239
 QY 73 PWSATVLOQTY-HAHRSDALQLGLKHNYCRNPNRRRPMVCYVQVGLKPLVQECVHDC 131
 DB 240 PWAS-----EATYRNVTAEQARNWGLGHAFCRNPDNDIRPWCFLNDRLSWEYCDLAQC 295
 QY 132 -----ADKKPSPEEL 144
 DB 296 QTPTQAAPTVPSPRLHVLMPAQPAPKPTTRTPPSQTPGALPAKREQPSLTRNG 355
 QY 145 KFCGQ---KTLRPRFKIIGGEFTTIENQFWFAIYRRHRGGSVTVYVCGSLISPCWVIS 201
 DB 356 PLSCGQRLRKSLSMTTRVVGVLALRGADHPYIAALYWGHS-----FCAGSLIAPCWILT 409
 QY 202 ATHCFIDYPKKEDYIVYLGSRSLNSNTQGMKFEVENLILHKOYSADTLAHNDIALKKI 261
 DB 410 AAHCLQDRPAPEDLTVVVLGQERRNHSCEPCQTALAVRSYRLHEAFS--PVSQYHDLALLRL 467
 QY 262 R-SKEGRCAQPSRTIOTICLPSMYNDPQFCTSCEITGFGKENSTDYLYPQLKMTVVKLI 320
 DB 468 QEDAGSCALLSPYQVPCLPSSGAARSETTLCQVAGWHQFEGAEYASFLQEAQPPFL 527
 QY 321 SHRECOQPHYGVSEVTTKMLCAADPQWKTDSCQDGGPLVCSLQ---RMTLTGVSWG 377
 DB 528 SLERCAPDVHGSSILPGMLCAGFLEGDTACQDGGPLVCEQDAERLTLQGIISWG 587
 QY 378 RGCALXDKPCVYTVRSHFLPWIRSHNT 403
 DB 588 SGCGRNKPVGVTDVAYYLAWIREHT 613
 RESULT 20
 FA12_BOVIN
 ID FA12_BOVIN STANDARD; PRT; 593 AA.
 AC P98140;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)
Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
(HAF) (Fragment).
F12.
Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=94242782; PubMed=8186251;
RA Shibuya Y., Semba U., Okebe H., Kambara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor
XII): comparison with human and guinea pig molecules.";
RL Biochim. Biophys. Acta 1206:63-70(1994).
[2]
SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77192112; PubMed=861210;
RA Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman
factor).";
RL Biochemistry 16:2270-2278(1977).
CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
VII to form factor VIIa and factor XI to form factor Xla.
CC -1- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HWM KININOGEN FORM A
COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS
CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/
KALLIKREIN CLEAVAGE SITE.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: S70164; AAB30804.2; -;
DR PIR: S45281; S45281.
DR KSSP: P00763; IDPO.
DR MEROPS: S01.211; -;
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000083; Fibrinctn1.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00039; fn1; 1.
DR Pfam: PF00040; fn2; 1.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00113; FNTYPEII.
DR PRINTS: PR00018; KRINGLE.
DR ProDom: PD000935; FN_Type_II; 1.
DR ProDom: PD000395; Kringle; 1.
SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01185; EGF_2; FALSE NEG.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1 1
FT SIGNAL <1 9 POTENTIAL.
FT CHAIN 10 349 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 350 593 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT DOMAIN 37 78 FIBRONECTIN TYPE-II.
FT DOMAIN 84 121 EGF-LIKE 1.
FT DOMAIN 123 163 FIBRONECTIN TYPE-I.
FT DOMAIN 164 200 EGF-LIKE 2.
FT DOMAIN 207 287 KRINGLE.
FT DOMAIN 297 333 PRO-RICH.
FT DOMAIN 350 593 SERINE PROTEASE.
FT ACT_SITE 389 389 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 438 438 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 541 541 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 88 100 BY SIMILARITY.
FT DISULFID 94 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 125 153 BY SIMILARITY.
FT DISULFID 151 160 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 173 188 BY SIMILARITY.
FT DISULFID 190 199 BY SIMILARITY.
FT DISULFID 207 287 BY SIMILARITY.
FT DISULFID 230 269 BY SIMILARITY.
FT DISULFID 258 282 BY SIMILARITY.
FT DISULFID 336 463 BY SIMILARITY.
FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 382 452 BY SIMILARITY.
FT DISULFID 413 416 BY SIMILARITY.
FT DISULFID 479 547 BY SIMILARITY.
FT DISULFID 510 526 BY SIMILARITY.
FT DISULFID 537 568 BY SIMILARITY.
FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
FT CARBOHYD 241 241 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;
Query Match 27.9%; Score 642; DB 1; Length 593;
Best Local Similarity 35.1%; Pred.No. 2.2e-43;
Matches 155; Conservative 60; Mismatches 167; Indels 60; Gaps 15;
QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFKGQHCIDKSKTCYE--GNGHFYRGKAS 63
Db 166 QVCRTNPCLNGDSCLQAE--GHRLCRCAPSPAGRLCDVDLKASCYDDRRDGLSYRMAG 222
QY 64 TDTMGPPCLPWNASATVLOQTY-HAHRSDALQLGLGHNYCRNPDRRRRWCVYVQGLKPL 122
Db 223 TTLSGAPCQSWAS----EATYWNVTAEQVLNGLGDHAFCRNPNDTRWCFTWKGDRLS 278
QY 123 VQBCMVHDC-ADG-----KKPSS-----PPELKF-----QCGQ 150
Db 279 WNYCLAPCAAAGHEHFFLPSPALQKPESTTQTPLSLTSGWCSPTPLASGGPGCGQ 338
QY 151 ---KTLRPFKLIIGGFTTIEQNPFAALYRRHGGSVTVVCGGSLISPCWVISATHCFI 207
Db 339 RLRLKWLSSLNRVVGGVLPGHVALPGHPIYALYWDQ-----HFCAGSLIAPCWILTAHCLQ 392

Qy 208 DYPKEDYVYVGRSLNNTQGENKFEVENLILHKOYSAOTLAHNDIALKIR-SKEG 266
 Db 393 NRPAPKELTVVLGQDRHNQSCQOTLAVRDLRHLHEAFPTYQH--DLALVRLQESADG 450
 Qy 267 RCAOPSRITQICLPKMYNDPQFGS--CELTGKGENSTDYLPQELKMTVVKLISHRE 324
 Db 451 CCAHPSFVQVCLPSTAAAESEAAVCEVAGWHQFEGE--YSSFLQEAQVPLIDPQR 509
 Qy 325 CQPHYGYSEVTTKMLCAADPQWKTDSCQDGGGGLVC--SLQGRMTLTGIVSWGRCGA 381
 Db 510 CSAFDVHGAAFTQGLMCLAGFLEGGTDACQDGGGGLVCEDETPERQLILRGIWMSGCG 569
 Qy 382 LKDKPGVYTVRSHLPWIRSH 403
 Db 570 NRLKPGVYTVANVLAIRREHT 591

RESULT 21
 PLMN_PIG STANDARD; PRT; 790 AA.
 AC P06867;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7).
 CN PLG.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP Schaller J., Marti T., Roessellet S.J., Kaempfer U., Rickli E.E.;
 RA "Amino acid sequence of the heavy chain of porcine plasmin. Comparison
 RT of the carbohydrate attachment sites with the human and bovine
 RT species.";
 RL Fibrinolysis 1:91-102(1987).
 RN [2]
 RP SEQUENCE OF 450-790.
 RN MEDLINE=85203907; PubMed=3846533;
 RA Marti T., Schaller J., Rickli E.E.;
 RT "Determination of the complete amino-acid sequence of porcine
 RT miniplasminogen.";
 RL Eur. J. Biochem. 149:279-285 (1985).
 RN [3]
 RP CARBOHYDRATE-LINKAGE SITES.
 RX MEDLINE=88185329; PubMed=3356193;
 RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
 RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
 RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
 RT plasminogen. Species specificity in relation to sialylation and
 RT fucosylation patterns.";
 RL Eur. J. Biochem. 173:57-63(1988).
 CC -|- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -|- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -|- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -|- PTM: N-LINKED GLYCAN CONTAINS N-ACETYLACTOSAMINE, SIALIC ACID AND
 CC IS CORE FUCOSYLATED. O-LINKED GLYCANS CONSIST OF GAL-GALNAc
 CC DISACCHARIDE WITH IS MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES
 CC (MICROHETEROGENEITY).
 CC -|- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -|- SIMILARITY: Contains 5 kringle domains.
 DR PIR: S03733; PLRG.
 DR HSP: P00747; SHPG.
 DR MEROPS: S01.233; --
 DR GlycoSuiteDB: P06867; --
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000001; Kringle.
 DR InterPro: IPR003014; PAN.
 DR InterPro: IPR003609; Pan_app.
 DR InterPro: IPR003966; Prothrombin.
 DR InterPro: IPR001254; Ser_protease_Try.
 DR Pfam: PF00051; kringle; 5.
 DR Pfam: PF00024; PAN; 1.
 DR Pfam: PF00089; tryptsin; 1.
 DR PRINTS: PR00722; CHYMOTRYPSIN.
 DR PRINTS: PR00018; KRINGLE.
 DR PRINTS: PR01505; PROTHROMBIN.
 DR PRODOM: PD000395; Kringle; 5.
 DR SMART: SM00130; KR; 5.
 DR SMART: SM00473; PAN AP; 1.
 DR SMART: SM00020; Tryp_SPC; 1.
 DR PROSITE: PS00021; KRINGLE_1; 5.
 DR PROSITE: PS00070; KRINGLE_2; 5.
 DR PROSITE: PS00240; TRYPSIN_DOM; 1.
 DR PROSITE: PS00134; TRYPSIN_HIS; 1.
 DR PROSITE: PS00135; TRYPSIN_SER; 1.
 DR Hydrolase: Serine protease; Plasma; Glycoprotein; Fibrinolysis;
 KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
 FT CHAIN 1 560
 FT PLASMIN HEAVY CHAIN A.
 FT CHAIN 561 790
 FT PLASMIN LIGHT CHAIN B.
 FT DOMAIN 561 790
 FT SERINE PROTEASE.
 FT DOMAIN 84 162
 FT KRINGLE 1.
 FT DOMAIN 166 243
 FT KRINGLE 2.
 FT DOMAIN 256 333
 FT KRINGLE 3.
 FT DOMAIN 358 435
 FT KRINGLE 4.
 FT DOMAIN 461 540
 FT KRINGLE 5.
 FT ACT SITE 602 602
 FT CHARGE RELAY SYSTEM.
 FT ACT SITE 645 645
 FT CHARGE RELAY SYSTEM.
 FT ACT SITE 740 740
 FT CHARGE RELAY SYSTEM.
 FT CARBOHYD 289 289
 FT N-LINKED (GLCNAC. . .).
 FT /FTID=CAR_000019.
 FT O-LINKED (GALNAC. . .).
 FT /FTID=CAR_000020.
 FT CARBOHYD 340 340
 FT O-LINKED (GALNAC. . .).
 FT /FTID=CAR_000020.
 SQ SEQUENCE 790 AA; 88592 MW; F04EA06E74BCD59E CRC64;
 Query Match 21.98; Score 504.5; DB 1; Length 790;
 Best Local Similarity 35.48; Fred. No. 2.3e-32;
 Matches 130; Conservative 45; Mismatches 141; Indels 51; Gaps 14;
 Qy 45 DSKTCVCEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHR---SDALQLGLGKH 100
 Db 456 DLSEDCMFGNGKRYGKRAITVAGVPCQEWAA-----QEPHRSITPTETNPRAGLEK-N 509
 Qy 101 YCRNPD-NRRRPNVCYQVGLKPLVQECMVHDCADGKKPSPBELKPCQCKTLRPR--- 156
 Db 510 YCRNPDGDDGPKCYT-TNPQKLFYCDVPQCVT-----SFDCKPKVPEKPCP 558
 Qy 157 FKIIIGEFTEIENQPFALYRHRGSSVYVCGGSLISPCWVISATHCFIDYPKKEDIY 216
 Db 559 ARVVGCVSIPHPWQISLRYRG-----HFCGGLISPEWVLTAKHLEKSSSPSSYK 614
 Qy 217 VYLGSRSLNSNTQEMKFEVENLILHKYSAADTLAHNDIALKIRSKRCQAPSRITQ 276
 Db 615 VILGAHEEYHLGSGVQIDVSKLF--KEPS-----EADIALKLSSP-----AVITDKVI 662
 Qy 277 TICLPWYNDPQF-----GTSCEITGKGENSTDYLPBOLKMTVVKLISHRECOQPHYYG 332
 Db 663 PACLPT-----PNYVADRTACTYTGWGETKGT--YGAGLLKEARLPVIEKNVCNRYEYL 716
 Qy 333 SEVTTKMLCAADPQWKTDSCQDGGGGLVCSLQGRMTLTGIVSWGRCALKQKPGYTVR 392
 Db 717 GKVSFNELCAGHLAGGIDSCQDGGGGLVCFEKDKYILQGVTSWGLGCALPNKPGYVVR 776

QY 393 SHELPMI 399
Db 777 SRFVTWI 783

RESULT 22
PLMN_MOUSE
ID PLMN_MOUSE STANDARD; PRT; 812 AA.
AC F20918; Q8CIS2; Q91WJ5;
DT 01-FEB-1991 (Rel. 17, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91184812; PubMed=2081600;
RA Degen S.J., Bell S.M., Schaefer L.A., Elliott R.W.;
RT "Characterization of the cDNA coding for mouse plasminogen and
RT localization of the gene to mouse chromosome 17";
RL Genomics 8:49-61(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic sequence analysis in the mouse t-complex region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC Tissue=Liver;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Richards S., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,
RA Villalón D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 1-16 FROM N.A.
RC STRAIN=129/SvJ; Tissue=Liver;
RX PubMed=12477932;
RA Bannach F.G., Gutierrez A., Fowler B.J., Bugge T.H., Degen J.L.,
RA Farmer R.J., Miles L.A.;
RT "Localization of regulatory elements mediating constitutive and
RT cytokine-stimulated plasminogen gene expression.";
RL J. Biol. Chem. 277:38579-38588(2002).
RN [5]
RP CHARACTERIZATION OF ANGIOSTATIN, AND PARTIAL SEQUENCE.
RX MEDLINE=95042728; PubMed=7525077;
RA O'Reilly M.S., Holmgren L., Shing Y., Chen C., Rosenthal R.A.,
RA Moses M., Lane W.S., Cao Y., Sage E.H., Folkman J.;
RT "Angiostatin: a novel angiogenesis inhibitor that mediates the
RT suppression of metastases by a Lewis lung carcinoma.";

Cell 79:315-328(1994).
-!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
LAMININ AND VON WILLEBRAND FACTOR.
-!- FUNCTION: ANGIOSTATIN IS AN ANGIOGENESIS INHIBITOR THAT BLOCKS
NEOVASCULARIZATION AND GROWTH OF EXPERIMENTAL PRIMARY AND
METASTATIC TUMORS IN VIVO.
-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
higher selectivity than trypsin. Converts fibrin into soluble
products.
-!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
-!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
-!- MISCELLANEOUS: IN THE PRESENCE OF THE INHIBITOR, THE ACTIVATION
INVOLVES ONLY CLEAVAGE AFTER ARG-581, RESULTING IN 2 CHAINS HELD
TOGETHER BY 2 DISULFIDE BONDS. WITHOUT THE INHIBITOR, THE
ACTIVATION INVOLVES ALSO REMOVAL OF THE ACTIVATION PEPTIDE.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
-!- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; J04766; AAA50168.1; -;
EMBL; AF481053; AAM22156.1; -;
EMBL; BC014773; AAH14773.1; -;
EMBL; AY134430; AAN15805.1; -;
PIR; A38514; PLMS.
HSSP; P00747; 1PMK.
MEROPS; S01.233; -;
MGD; MGI:97620; Pig.
GO; GO:0016506; P:Apoptosis activator activity; IDA.
GO; GO:0006915; P:Apoptosis; IDA.
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR000001; Kringle.
InterPro; IPR003014; PAN.
InterPro; IPR003609; Pan_app.
InterPro; IPR003966; Prothrombin.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00051; kringle; 5.
Pfam; PF00024; PAN; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
PRINTS; PR01505; PROTHROMBIN.
ProDom; PD000395; Kringle; 5.
SMART; SM00130; KR; 5.
SMART; SM00473; PAN_AP; 1.
SMART; SM00020; TRYD_SPE; 1.
PROSITE; PS00021; KRINGLE_1; 5.
PROSITE; PS00070; KRINGLE_2; 5.
PROSITE; PS00240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase. Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW Signal.
FT CHAIN 1 19 PLASMINOGEN.
FT CHAIN 20 812 PLASMIN HEAVY CHAIN A.
FT CHAIN 20 581 ACTIVATION PEPTIDE.
FT PEPTIDE 20 97


```
FT CHAIN 98 581 PLASMIN SHORT FORM OF CHAIN A.
FT CHAIN 98 7436 ANGIOTATIN.
FT CHAIN 582 812 PLASMIN LIGHT CHAIN B.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 184 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 377 454 KRINGLE 4.
FT DOMAIN 481 560 KRINGLE 5.
FT DOMAIN 582 812 SERINE PROTEASE.
FT ACT_SITE 667 667 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 762 762 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 49 73 BY SIMILARITY.
FT DISULFID 53 61 BY SIMILARITY.
FT DISULFID 103 181 BY SIMILARITY.
FT DISULFID 124 164 BY SIMILARITY.
FT DISULFID 152 176 BY SIMILARITY.
FT DISULFID 185 262 BY SIMILARITY.
FT DISULFID 188 316 BY SIMILARITY.
FT DISULFID 206 245 BY SIMILARITY.
FT DISULFID 234 257 BY SIMILARITY.
FT DISULFID 275 352 BY SIMILARITY.
FT DISULFID 324 347 BY SIMILARITY.
FT DISULFID 377 454 BY SIMILARITY.
FT DISULFID 398 437 BY SIMILARITY.
FT DISULFID 426 449 BY SIMILARITY.
FT DISULFID 481 560 BY SIMILARITY.
FT DISULFID 502 543 BY SIMILARITY.
FT DISULFID 531 555 BY SIMILARITY.
FT DISULFID 568 687 INTERCHAIN (BY SIMILARITY).
FT DISULFID 578 586 INTERCHAIN (BY SIMILARITY).
FT DISULFID 609 625 BY SIMILARITY.
FT DISULFID 701 768 BY SIMILARITY.
FT DISULFID 731 747 BY SIMILARITY.
FT DISULFID 758 786 BY SIMILARITY.
FT CONFLICT 235 235 R -> H (IN REF. 1).
FT CONFLICT 525 525 G -> D (IN REF. 1).
FT CONFLICT 649 649 S -> L (IN REF. 1).
FT CONFLICT 649 649 S -> L (IN REF. 1).
SQ SEQUENCE 812 AA; 90781 MW; 2417326056A2FFD2 CRC64;

Query Match
Best Local Similarity 32.5%; Score 501; DB 1; Length 812;
Matches 135; Conservative 48; Mismatches 170; Indels 62; Gaps 15;

Qy 10 NCDCLNGTGVSNKYFNHWCNPK--KFGQHQEI-----DKSKTYEGNGH 56
Db 428 NPDGDKGPWCYTTDPVSRWEYCNLRCSETGSGVVELFTVSGPSPGSDSETDCMYGNK 487
Qy 57 FYRKASDTMGRPLPWN SATVLQQTVAHRS DALQ----LGLGKKNYCNPD-NRRRP 111
Db 488 DYRGKTAVTAAGTFCQGWAA-----QPHRHSIFPTQTNFAGLEK-NYCRNPGDVNGP 541
Qy 112 WCYVQVGLKPLVQCNVHDCADGKKPSPPBELKFCQCKTLRPR---FKIIGGEFTTIE 168
Db 542 WCYT-TNPRKLYDCDIPLCASAS-----SFEGCKPQVEPKPCGRVVGCVANPH 591
Qy 169 NQPFALYRHRGSGVYVCGSLSPCWVISATHCFIDYPKKEDYVILGRSLNSNT 228
Db 592 SWPQISLIRTRFTG---QHFGGGLTIAPEWVLTAAHCLKSSRPFYKVLGAHEEYIRG 648
Qy 229 QGEMKFEVENLHKDYASDTLAHNDIALKIRSKRCRCAQPSRTQICLPSMYNDPQ 288
Db 649 SDVQEISVAKLLE-----PNNRDIALKL-----SRPATITDKVIPACLPS----PN 692
Qy 289 F-----GTSCEITGFGKSNSTLYLPEQLKMTVVKLISHRECOQPHYGVSEVTTKWLCAAD 344
Db 693 YMWADRTICITYTNGETQGT---FGAGRLKEAQLPVIENKVCNRVLYNNRNVKSTELCAGQ 750
Qy 345 PQWKTDCQDGGGLVCSLQGRMTLTGIVSGWGGCALKOKPGYVTVRSHPLPW 399
Db 751 LAGVDSQDGGGLVCFEAKDKVILQGVTSWGLGRCARPNKPGYVTVRSHPLPW 805
```

```
RESULT 23
PLMN HUMAN
ID PLMN HUMAN STANDARD; PRT; 810 AA.
AC P00747;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7) [Contains: Angiostatin].
GN PLG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90202879; PubMed=2318848;
RA Petersen T.E., Martzen M.R., Ichinose A., Davie E.W.;
RT "Characterization of the gene for human plasminogen, a key proenzyme
in the fibrinolytic system.";
RL J. Biol. Chem. 265:6104-6111(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87162490; PubMed=3030813;
RA Forsgren M., Raden B., Israelsson M., Larsson K., Heden L.-O.;
RT "Molecular cloning and characterization of a full-length cDNA clone
for human plasminogen.";
RL FEBS Lett. 213:254-260(1987).
RN [3]
RP SEQUENCE OF 20-810.
RA Sottrup-Jensen L., Petersen T.E., Magnusson S.;
RL Submitted (JUL-1977) to the FIR data bank.
RN [4]
RP SEQUENCE OF 292-810 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [5]
RP SEQUENCE OF 20-100.
RX MEDLINE=75093329; PubMed=122932;
RA Wiman B., Wallen P.;
RT "Structural relationship between 'glutamic acid' and 'lysine' forms
of human plasminogen and their interaction with the NH2-terminal
activation peptide as studied by affinity chromatography.";
RL Eur. J. Biochem. 50:489-494(1975).
RN [6]
RP SEQUENCE OF 95-580; 581-626; 657-700 AND 732-810.
RA Sottrup-Jensen L., Claeys H., Zajdel M., Petersen T.E., Magnusson S.;
RL (In) Davidson J.F., Rowan R.M., Samama M.M., Desnoyers P.C. (eds.);
RT Progress in chemical fibrinolysis and thrombolysis, pp.3:191-209,
Raven Press, New York (1978).
RN [7]
RP SEQUENCE OF 483-604.
RX MEDLINE=76043692; PubMed=126863;
RA Wiman B., Wallen P.;
RT "Amino-acid sequence of the cyanogen-bromide fragment from human
plasminogen that forms the linkage between the plasmin chains.";
RL Eur. J. Biochem. 58:539-547(1975).
RN [8]
RP SEQUENCE OF 581-810.
RX MEDLINE=7725245; PubMed=142009;
RA Wiman B.;
RT "Primary structure of the B-chain of human plasmin.";
RL Eur. J. Biochem. 76:129-137(1977).
RN [9]
RP ACTIVE SITE.
RX MEDLINE=73149248; PubMed=4694729;
RA Robbins K.C., Bernabe P., Arzadon L., Summaria L.;
RT "The primary structure of human plasminogen. II. The histidine loop
of human plasmin: light (B) chain active center histidine sequence.";
RL J. Biol. Chem. 248:1631-1633(1973).
```

RN X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 RX MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringle 1 domain of human
 plasminogen in complexes with the ligands epsilon-aminocaproic acid
 and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
 RL Biochemistry 35:2567-2576(1996).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
 RX MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringle 1 domain of human
 plasminogen in complexes with the ligands epsilon-aminocaproic acid
 and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
 RL Biochemistry 35:2567-2576(1996).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
 RX MEDLINE=98198034; PubMed=9521645;
 RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
 RA Castellino F.J.;
 RT "Structure and ligand binding determinants of the recombinant kringle
 5 domain of human plasminogen.";
 RL Biochemistry 37:3238-3271(1998).
 RN [23]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237157; PubMed=8181475;
 RA Rejante M.R., Llinas M.;
 RT "1H-NMR assignments and secondary structure of human plasminogen
 kringle 1.";
 RL Eur. J. Biochem. 221:927-937(1994).
 RN [24]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237158; PubMed=8181476;
 RA Rejante M.R., Llinas M.;
 RT "Solution structure of the epsilon-aminohexanoic acid complex of
 human plasminogen kringle 1.";
 RL Eur. J. Biochem. 221:939-949(1994).
 RN [25]
 RP STRUCTURE BY NMR OF 183-354.
 RX MEDLINE=96194156; PubMed=8652577;
 RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
 RA Rickli E.E.;
 RT "Recombinant gene expression and 1H NMR characteristics of the
 kringle (2 + 3) supermodule: spectroscopic/functional individuality
 of plasminogen kringle domains.";
 RL Biochemistry 35:2357-2364(1996).
 RN [26]
 RP STRUCTURE BY NMR OF 374-461.
 RX MEDLINE=90219023; PubMed=2157850;
 RA Atkinson R.A., Williams R.J.P.;
 RT "Solution structure of the kringle 4 domain from human plasminogen by
 1H nuclear magnetic resonance spectroscopy and distance geometry.";
 RL J. Mol. Biol. 212:541-552(1990).
 RN [27]
 RP VARIANTS PHE-374 AND THR-620.
 RN [27]

RT "The refined structure of the epsilon-aminocaproic acid complex of
 human plasminogen kringle 4.";
 RL Biochemistry 30:10589-10594(1991).
 RN [20]
 RP X-RAY CRYSTALLOGRAPHY (1.67 ANGSTROMS) OF 376-454.
 RX MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringle 1 domain of human
 plasminogen in complexes with the ligands epsilon-aminocaproic acid
 and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
 RL Biochemistry 35:2567-2576(1996).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 102-181.
 RX MEDLINE=96180681; PubMed=8611560;
 RA Mathews I.I., Vanderhoff-Hanaver P., Castellino F.J., Tulinsky A.;
 RT "Crystal structures of the recombinant kringle 1 domain of human
 plasminogen in complexes with the ligands epsilon-aminocaproic acid
 and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid.";
 RL Biochemistry 35:2567-2576(1996).
 RN [22]
 RP X-RAY CRYSTALLOGRAPHY (1.66 ANGSTROMS) OF 480-563.
 RX MEDLINE=98198034; PubMed=9521645;
 RA Chang Y., Mochalkin I., McCance S.G., Cheng B., Tulinsky A.,
 RA Castellino F.J.;
 RT "Structure and ligand binding determinants of the recombinant kringle
 5 domain of human plasminogen.";
 RL Biochemistry 37:3238-3271(1998).
 RN [23]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237157; PubMed=8181475;
 RA Rejante M.R., Llinas M.;
 RT "1H-NMR assignments and secondary structure of human plasminogen
 kringle 1.";
 RL Eur. J. Biochem. 221:927-937(1994).
 RN [24]
 RP STRUCTURE BY NMR OF 96-184.
 RX MEDLINE=94237158; PubMed=8181476;
 RA Rejante M.R., Llinas M.;
 RT "Solution structure of the epsilon-aminohexanoic acid complex of
 human plasminogen kringle 1.";
 RL Eur. J. Biochem. 221:939-949(1994).
 RN [25]
 RP STRUCTURE BY NMR OF 183-354.
 RX MEDLINE=96194156; PubMed=8652577;
 RA Soehndel S., Hu C.-K., Marti D., Affolter M., Schaller J., Llinas M.,
 RA Rickli E.E.;
 RT "Recombinant gene expression and 1H NMR characteristics of the
 kringle (2 + 3) supermodule: spectroscopic/functional individuality
 of plasminogen kringle domains.";
 RL Biochemistry 35:2357-2364(1996).
 RN [26]
 RP STRUCTURE BY NMR OF 374-461.
 RX MEDLINE=90219023; PubMed=2157850;
 RA Atkinson R.A., Williams R.J.P.;
 RT "Solution structure of the kringle 4 domain from human plasminogen by
 1H nuclear magnetic resonance spectroscopy and distance geometry.";
 RL J. Mol. Biol. 212:541-552(1990).
 RN [27]
 RP VARIANTS PHE-374 AND THR-620.
 RN [27]

Query Match 21.7%; Score 499; DB 1; Length 810;
 Best Local Similarity 33.5%; Pred. No. 6.4e-32;
 Matches 139; Conservative 44; Mismatches 168; Indels 64; Gaps 16;
 10 NCDLNGGTCVSNKYFSNIHWCNCPKFGGQ-----HCEIDKSKTCYEGNGH 56
 428 NPDADKGPWCFTTDPVSWEYCNLKKSCGTEASVAPPVLLPDVETSEEDCFEGNCK 487
 57 FYRGKASTDTMCRPCLPWN SATVLOQTVAHR-----SDALQLGLGHNYCRNPD-NRRRP 111
 488 GYRGKRAITVTGTPCQDWA-----QEPHRHSIFTFETNFRAGLEK-NYCRNPDGVDGP 541
 112 WCYVQVGLKPLVQECQWHD CADKPKSPPELKFQCGQKTLRPR---FKITGGFTTIE 168
 542 WCYT-TNPRKLYDYCDVPOCA---APS-----FDCGKPQVEPKKCPGRVVGCVVAHPH 590

Query Match 21.6%; Score 497; DB 1; Length 810;
Best Local Similarity 33.7%; Pred. No. 9.2e-32;
Matches 140; Conservative 46; Mismatches 165; Indels 64; Gaps 16;

QY 10 NCCLNGGTCVSNKYFSNIHWCNCPKFKGQ-----HCEIDSKSKTCYEGNGH 56
Db 428 NPADKGPWCFTTDPVSRWEYCNLKCSTGEGVAAPPPVAQLPDAETSEEDCMFGNGK 487

QY 57 FYRGKASTDTMGRPCLPWN SATVLQOTYHAHR-----SDALQLGLGKHNYCRNP 111
Db 488 GYRGKATTVTGTPCQEWAA-----QEPHSHRIFTPTNPAGLEK-NYCRNP DGVGGP 541

QY 112 WCYVQVGLKPLVQECMVHDCADGKSPPEELKFCGQKTLRPR---FKIIGETTIE 168
Db 542 WCYT-TNPKRLFDYCDVPOCAAS-----SFDGKRPQVPEPKPCGRVGGCVAYPH 590

QY 169 NOPWFAAIYRRHRGGSVTYVCGGSLSPCWVISATHCFIDYPKKEDIYVVLGRSLNSNT 228
Db 591 SWPW--QISLRTELG--MHFCGGTLISPEWVLTAAHCLSKSRPSFYKVLGAHR-----641

QY 229 QGEMKEVENLILHKDYSADTLAHNDIALLKIRSKGRCQPSRTIQTICLPSMYNDPQ 288
Db 642 EVHLEPHVQEIYVSKMFSEPARA---DIALKLSSP---AIIIDKVIAPCLPS-----DN 690

QY 289 F-----GTSCEITGFGKENSTDYLPQLKMTVVKLISHRECOPHYVGSVTTKMLCAAD 344
Db 691 YVVADETECFITGWGTQGT--YGAGLLKEARLPVIENKVCNRYEFLNGVTVKITELCAGH 748

QY 345 POWKTDSCQDGGGLPVCSLQGRMTLTGIVSMRGCAKDKPGVYTRVSHFLPWI 399
Db 749 LAGGTDSQDGGGLPVCFEKDKYILQGVTSWGLGCARPKNKPGVYVRSRVFTWI 803

RESULT 25
PLMN_CANFA STANDARD; PRT; 333 AA.

AC P80009;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Plasminogen (BC 3.4.21.7) (Fragment).
GN PLG.

OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.

RC TISSUE=Plasma;
RX MEDLINE=90175323; PubMed=2626424;
RA Schaller J., Straub C., Kaempfer U., Rickli B.E.;
RL "Complete amino acid sequence of canine miniplasminogen.";
RL Protein Seq. Data Anal. 2:445-450(1989).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UKONINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. ACTIVATED WITH UKONINASE AND HIGH CONCENTRATIONS OF
CC STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains at least 1 kringle domain.

HSP; P00747; SHP.
MEROPS; S01.233; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003966; Prothrombin.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen.
FT NON_TER 1 1
FT CHAIN <1 103 PLASMIN HEAVY CHAIN A.
FT CHAIN 104 333 PLASMIN LIGHT CHAIN B.
FT DOMAIN 4 333 KRINGLE 5.
FT DOMAIN 104 333 SERINE PROTEASE.
FT DISULFID 4 83 BY SIMILARITY.
FT DISULFID 25 66 BY SIMILARITY.
FT DISULFID 54 78 BY SIMILARITY.
FT DISULFID 90 208 INTERCHAIN (BY SIMILARITY).
FT DISULFID 100 108 INTERCHAIN (BY SIMILARITY).
FT DISULFID 130 146 BY SIMILARITY.
FT DISULFID 222 289 BY SIMILARITY.
FT DISULFID 252 268 BY SIMILARITY.
FT DISULFID 279 307 BY SIMILARITY.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 188 188 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 283 283 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 152 152 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 186 186 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 264 264 STREPTOKINASE-BINDING SITE (PROBABLE).
FT SITE 277 277 SITE OF SUBSTRATE SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 333 AA; 36678 MW; C8C0271B6C6AC8D4 CRC64;

Query Match 21.5%; Score 494.5; DB 1; Length 333;
Best Local Similarity 34.8%; Pred. No. 5.3e-32;
Matches 126; Conservative 49; Mismatches 136; Indels 51; Gaps 14;

QY 50 CYEGNGHYRGKASTDTMGRPCLPWN SATVLQOTYHAHR-----SDALQLGLGKHNYCRNP 105
Db 4 CMFGNGKGYRGKATTVMGIPQEWAA-----QEPHHSIFTPTNPQAGLEK-NYCRNP 57

QY 106 D-NRRRPPWCYCVQGLKPLVQECMVHDCADGKSPPEELKFCGQKTLRPR---FKIIG 161
Db 58 DGVNNGPWCYT-MNQRKLFYCDVQCV-----STSDCKRPQVPEPKPCGRVVG 106

QY 162 GEFTTIENQPFPAALYRRHRGGSVTYVCGGSLSPCWVISATHCFIDYPKKEDIYVILGR 221
Db 107 GCVANPHSPWQISLRTRYG---KHFCGGTLISPEWVLTAAHCLSKSRPSYKVLG- 161

QY 222 SELNSNTGEMKFEVENLILHKDYSADTLAHNDIALLKIRSKGRCQPSRTIQTICLP 281
Db 162 ---AHKEVNLESVDQIEVYKLFLEPTRA---DIALKLSSP---AVITSKVIPACLIP 210

QY 282 SWYNDPQ---GTSCEITGFGKENSTDYLPQLKMTVVKLISHRECQPHYVSEVTT 337
Db 211 ---PPNVTVADRTLCTYITGNETQGT---YGAGLLKEARLPVIENKVCNRYEFLNGRYS 264

QY 338 KMLCAADPQWKTDSCQDGGGLPVCSLQGRMTLTGIVSMRGCAKDKPGVYTRVSHFLP 397
Db 265 TELCAGNLAGTDSQDGGGLPVCFEKDKYILQGVTSWGLGCARPKNKPGVYVRSRVFT 324

```
QY 398 WI 399
  ||
Db 325 WI 326

RESULT 26
ID PLMN SHEEP STANDARD; PRT; 343 AA.
AC P81286;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen (EC 3.4.21.7) (Fragment).
GN PLG.
OS Ovis aries (sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP MEDLINE=93149995; PubMed=1492092;
RX Schaller J., Straub C., Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of ovine miniplasminogen.";
RL Protein Seq. Data Anal. 5:21-25(1992).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains at least 2 kringle domains.
DR PIR: B61545; B61545.
DR HSP: P00747; SHPG.
DR MEROPS: S01.233; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003966; Prothrombin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00051; kringle_1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR0018; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat.
FT NON TER 1
FT DOMAIN <1 140 HEAVY CHAIN A.
FT DOMAIN 141 >343 LIGHT CHAIN A.
FT DOMAIN <1 17 KRINGLE 4.
FT DOMAIN 41 120 KRINGLE 5.
FT DOMAIN 114 341 SERINE PROTEASE.
FT ACT_SITE 181 181 CHARGE RELAY SYSTEM.
FT ACT_SITE 224 224 CHARGE RELAY SYSTEM.

FT ACT_SITE 319 319 CHARGE RELAY SYSTEM.
FT NON TER 343 343
SQ SEQUENCE 343 AA; 37662 MW; 8DF6EBA92D596BE0 CRC64;

Query Match 21.4%; Score 493; DB 1; Length 343;
Best local similarity 35.1%; Pred. No. 7.2e-32;
Matches 127; Conservative 41; Mismatches 142; Indels 52; Gaps 13;

QY 50 CYEGNGHYFGKASTDTMGPCPLPWN SATVLQOYTHAH---RSDALQLGLKHNCRNP 105
  |||
Db 15 CMIGIGKGYGKKATTVAGVPCQEWAA-----QEPHRRGIFTETNFRAGLEK-NYCRNP 68
  |||
QY 106 D-NRRRPWCYVQVLKPLVQECMVHDCADGKPKSPPEELKFCQCGQKTLRPR---FKIIG 161
  |||
Db 69 DGDVNGPWCYT-TNPRKLFDCIPQC-----ESSFDCGKPKVKKPCARVVG 116
  |||
QY 162 GERTTIENQWFAAIRRHGGSVTVYVCGSLSPCWVISATHCFDIYPKKEDIVYVLR 221
  |||
Db 117 GCVATPHSWPQVSLRRSR-----EHFCGGTLLSPWVLTAAHCLDSILGPSFYVILGA 172
  |||
QY 222 SRLNSNTQGMKFEVENLIILHKDYSADTLAHHNDIALLKIRSKRGCAQPSRTIQTICLP 281
  |||
Db 173 HYEMAREASQVEIPVSEFLFLEPSRA-----DIALKLSSP-----AVITDEVIPACLP 220
  |||
QY 282 SMYNDPOF-----GTSCEITGFGKENSTDYLYPOLKMTVVKLISHRECQPHYYGSEVTT 337
  |||
Db 221 S-----PNYVVADKTVCVITGWTGTGT--FGVGRLEKARLPVIENKVCNRYEYLNGRVKS 274
  |||
QY 338 KMLCADPQWKTSQCGDSGGPLVCSLOGRMTLTGIVSWGCGALKDKPGVTVVSHFLP 397
  |||
Db 275 TELCAGLAGTSCQDSGGPLVCFEKKYIILQGVTSWGLGCAKPRKPGVYVRVSTVVP 334
  |||
QY 398 WI 399
  ||
Db 335 WI 336

RESULT 27
ID PLMN BOVIN STANDARD; PRT; 812 AA.
AC P06858; Q28162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Berglund L., Andersen M.D., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen cDNA.";
RL Int. Dairy J. 5:593-603(1995).
RN [2]
RN SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=85203906; PubMed=3846532;
RX Schaller J., Moser P.W., Danneegger-Muller G.A.K., Rosselet S.J.,
RA Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of bovine plasminogen. Comparison with
RT human plasminogen.";
RL Eur. J. Biochem. 149:267-278(1985).
RN [3]
RN SEQUENCE OF 706-812 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
RT human and bovine plasminogen.";
RL Biochemistry 23:4243-4250(1984).
RN [4]
```


RESULT 30

NETR HUMAN STANDARD; PRT; 875 AA.
 AC P56730; Q9UP16;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurotysin precursor (EC 3.4.21.-) (Motopsin) (Leydin).
 GN PRSS12.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98201705; PubMed=9540828;
 RA Proba K., Gschwend T.P., Sonderegger P.;
 RT "Cloning and sequencing of the cDNA encoding human neurotysin.";
 RL Biochim. Biophys. Acta 1396:143-147(1998).
 RN [2]
 RP SEQUENCE OF 615-875 FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=99203523; PubMed=10103056;
 RA Poorafshar M., Hellman L.;
 RT "Cloning and structural analysis of leydin, a novel human serine
 protease expressed by the Leydig cells of the testis.";
 RL Eur. J. Biochem. 261:244-250(1999).
 CC -!- FUNCTION: PLAYS A ROLE IN NEURONAL PLASTICITY AND THE PROTEOLYTIC
 ACTION MAY SUBSERVE STRUCTURAL REORGANIZATIONS ASSOCIATED WITH
 LEARNING AND MEMORY OPERATIONS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Brain and Leydig cells of the testis.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 4 SRCR domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 CC EMBL: AJ001531; GAA04816.1; -;
 CC EMBL: AR077298; AAD25919.1; -;
 CC HSSP: P00763; IDPO.
 CC Genew; HGNC:9477; PRSS12.
 CC MIM; 606709; -;
 CC MEROPS; S01.237; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.
 CC InterPro; IPR001190; Srcr_receptor.
 CC Pfam; PF00051; kringle; 1.
 CC Pfam; PF00530; SRCR; 4.
 CC Pfam; PF00089; trypsin; 1.
 CC PRINTS; PR00722; CHYMOTRYPSIN.
 CC PRINTS; PR00256; SPERACTRCPTR.
 CC Prodom; PD000395; kringle; 1.
 CC SMART; SM00130; KR; 1.
 CC SMART; SM00202; SR; 4.
 CC SMART; SM00020; Tryp_Spc; 1.
 CC PROSITE; PS00021; KRINGLE_1; 1.
 CC PROSITE; PS50070; KRINGLE_2; 1.
 CC PROSITE; PS00420; SRCR_1; 3.
 CC PROSITE; PS50287; SRCR_2; 4.
 CC PROSITE; PS50240; TRYPSIN_DOM; 1.
 CC PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Kringle; Repeat; Signal.

FT	SIGNAL	1	20	POTENTIAL.
FT	CHAIN	21	875	NEUROTYPYPSIN.
FT	DOMAIN	23	92	PRO-RICH.
FT	DOMAIN	93	165	KRINGLE.
FT	DOMAIN	170	271	SRCR 1.
FT	DOMAIN	280	381	SRCR 2.
FT	DOMAIN	387	487	SRCR 3.
FT	DOMAIN	500	601	SRCR 4.
FT	DOMAIN	619	875	SERINE PROTEASE.
FT	DOMAIN	619	875	ZYMOGEN ACTIVATION REGION.
FT	ACT SITE	630	631	REACTIVE BOND (POTENTIAL).
FT	ACT SITE	676	676	CHARGE RELAY SYSTEM.
FT	ACT SITE	726	726	CHARGE RELAY SYSTEM.
FT	ACT SITE	825	825	CHARGE RELAY SYSTEM.
FT	DISULFID	619	750	POTENTIAL.
FT	CARBOHYD	26	26	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	683	683	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	663	663	A -> V (IN REF. 2).
FT	CONFLICT	701	701	E -> V (IN REF. 2).
FT	CONFLICT	839	841	VVY -> AAL (IN REF. 2).
SQ	SEQUENCE	875 AA; 97011 MW; B66EC946DC208DC8	CRG64;	

Query Match

20.1%; Score 462; DB 1; Length 875;

Best Local Similarity 29.8%; Pred. No. 5.9e-29;

Matches 130; Conservative 60; Mismatches 156; Indels 90; Gaps 16;

QY	33	CPKFGGQH-C--EIDKSKTCYEG-----	-----NGHF----	57
DB	465	CSRQWGRHDCSHREDVSIACYPGGEGHRLSLGFPVRLMDGENKKEGRVEVFINGQWTI	-----	524
QY	58	-----YRGKASTDTM-----	-----	96
DB	525	CDDGWTDKDAVICQLGYKGFARATMAYFEGKGIHVDNVKCTGNERSLADCIQDI	-----	584
QY	97	GKHNYCRNPDRRRPWCYVQGLKPLVQECMVHDCADGKKPS--SPPEELKFCQGOKTLR	-----	154
DB	585	GRHN-CRHSEDAGVICDYF-----	-----GKKASGNSKESLSVCGRLRLH	625
QY	155	PRFK-IIGGEFTIENQPFPAIYRRHROGSVTVYVCGSLISPCWVISATHCFIDYPKK-	-----	212
DB	626	RRQKRIIGKNSLRGGWPQVSLRLKSHGDGLLCGATLLSSCWVLTAAHCFKRYGNST	-----	685
QY	213	EDYIVVLGRSLNSNTQGMKFEVENLIILHKDYSADTLAHNDIALLKIRSKRGCAQPS	-----	272
DB	686	RYAVRVGYHTLVPEEFEEIGVQVIVHREYRDRSDY--DIALVRLQGPEEQCARFS	-----	743
QY	273	RTIQTICLPSMYNDPQ-FQTSCEITFGKENSTDYLYPEQLKMTVVKLIHRECCQPHY	-----	331
DB	744	SHVLPACLPLWRERPOKTASNCVITGWG--DTGRAYSRTLQQAIPLLPKRFCEE--RY	-----	798
QY	332	GSEVTKMLCAAD--PQWKTDSQGDSCGPIVCSLQGR-MTLTGIVSWGRGCALKDKPGV	-----	388
DB	799	KGRFTGRMLCAGNLHEHKRVDSQGDSCGPIVCSLQGR-MTLTGIVSWGRGCALKDKPGV	-----	858
QY	389	YTRVSHFLPWIRSHTK 404		
DB	859	YTKVSFAFVPIKSVTK 874		

Search completed: December 3, 2003, 14:40:06

Job time : 15.2061 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 50.5697 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-3
Perfect score: 2301
Sequence: 1 SNELHQVSNCDLNGTCV.....VSHFLPWIRSHYKENGSL 411

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 23.*

- 1: sp_archea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacterioph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1923	83.6	433	6	Q8MIL0
2	1918	83.4	433	6	Q8MHI7
3	982.5	42.7	214	6	Q9XTI70
4	858.5	37.3	516	4	Q9BU99
5	842.5	36.6	231	11	Q8C6L2
6	826.5	35.9	559	11	Q91VP2
7	819	35.6	564	6	Q8MKB1
8	816.5	35.5	562	6	Q8SQ23
9	784.5	34.1	395	4	Q9BZW1
10	778	33.8	154	4	Q96SE8
11	726	31.6	653	11	Q8VCS4
12	693.5	30.1	616	6	Q97507
13	689	29.9	615	4	Q81Z25
14	677	29.4	157	6	Q91VA8
15	677	29.4	597	11	O35727
16	674.5	29.3	517	11	Q8K0D2

17	661.5	28.7	560	4	Q14520
18	585	25.4	128	6	O97587
19	501	21.8	501	11	Q91WJ5
20	498	21.6	810	4	Q15146
21	496	21.6	454	6	O46506
22	496	21.6	812	11	Q9R0W3
23	495.5	21.5	300	4	O96EF3
24	495	21.5	103	6	Q95M89
25	491.5	21.4	429	13	Q8AVB0
26	482.5	21.0	334	6	O46507
27	472.5	20.5	806	6	O18783
28	463.5	20.1	868	5	O9Y1V3
29	442.5	19.2	327	4	O8N171
30	435.5	18.9	284	4	Q8NF86
31	433.5	18.8	761	11	Q99JCB
32	426.5	18.5	267	5	Q9BK47
33	425.5	18.5	505	5	Q966V4
34	424	18.4	537	4	Q9BYE1
35	420	18.3	471	11	Q8CFE0
36	420	18.3	581	4	Q9BYE2
37	420	18.3	802	4	O81UE2
38	420	18.3	811	4	O8IU80
39	419.5	18.2	799	11	Q9DBI0
40	414	18.0	638	11	Q8R0P5
41	411.5	17.9	276	11	O8CGR6
42	409	17.8	371	11	Q8CJL6
43	409	17.8	445	11	Q8CJL7
44	405.5	17.6	422	4	Q8WVC1
45	402	17.5	455	11	Q8CDR0

ALIGNMENTS

RESULT 1
Q8MIL0
ID Q8MIL0 PRELIMINARY; PRT; 433 AA.
AC Q8MIL0;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY122285; AAM83187.1; -!
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.


```

DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_Spc; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match
Best Local Similarity 42.7%; Score 982.5; DB 6; Length 214;
Matches 178; Conservative 14; Mismatches 22; Indels 1; Gaps 1;

QY 67 MGRPCLPWSNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLKPLVQEC 126
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MDRPCLAWNSANVLTQTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVQVGLKPLVQEC 60

QY 127 MVHDCADGKPSPPBELKFCQCKTLRPFKLIIGGFTTIENQPWFAAIYRHRGGSVT 186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 KVHD-SSGKKPALPPQKLEFQCCQKALRPFKLIIGGFTTIENQPWFAAIYRHRGGSVT 119

QY 187 YVCGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNTQGMKFEVENILHKDYS 246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
120 YVCGSLISPCWVVSATHCFINQKXEDIYVILGRPLNSMTQGMKFEVEQIILHEGYR 179

QY 247 ADTLAHNDIALKIRSKRGCRQAPRTTICLP 281
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
180 ADTLAHNDIALKILSNNGQCAQPSRSIQICLP 214

RESULT 4
Q9BU99 PRELIMINARY; PRT; 516 AA.
AC Q9BU99;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC02795; AAH02795.1; -.
DR HSSP; P00750; IASH.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00051; Kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.

DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 516 AA; 57370 MW; BAB31901FDC96800 CRC64;

Query Match
Best Local Similarity 37.3%; Score 858.5; DB 4; Length 516;
Matches 187; Conservative 55; Mismatches 159; Indels 95; Gaps 10;

QY 1 SNELH-----QVPSNCDCLNGGTCVSNKYFYSNIHWCNCPKFGGQCHCEIDKSKTC 50
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
23 SQIHARFRGARSYQCCSEPRCFNGGTCQOALYFSDF-VQCPEGPAGKCCCEIDTATC 81

QY 51 YENGHFYRKASTDTWGRCLPWSNATVLOQTYHAHRSDALQLGLGKHNCRNPNRRR 110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
82 YEDQGI-SYRGITWSTAESGAECTWNSSALAKQKPSGRRPDAIRLGLGNHNYCRNPD RSK 141

QY 111 PMCYVQVGLKPLVQECMVHDCADG----- 134
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
142 PWCYVFKAGKYSSEFCSTPACSEGNSDCYFGNGSAYRGTHSLTESGASCLPWNMILIGK 201

QY 135 ----KKPS-----SPPEELK-----FQCG-QKLRP 155
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
202 VYTAQNPQAALGLGKHNCRNPDGDAKPMCHVKRRLTWECVDPSCSTGLRQYSQP 261

QY 156 RPKIIGGFTTIENQPWFAAIYRHR-RGGSVTVVCGSLISPCWVISATHCFIDYPKKED 214
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
262 QFRIKGLFADIASHPWQAIAFAKRRSPGERFLCGGILISSCWILLSAAHCFQERFPFPH 321

QY 215 YIVLGRSLNSNTQGMKFEVENILHKDYSADTLAHNDIALKIRSKRGCRQAPSR 274
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
322 LTVLGRITRVVPGEBEQPEKVIYVHKFDDDT--YNDIALQLKSDSSRCAQESSV 379

QY 275 IOTICLPNNDPQFGTSCBITGFGKENSDIYLPQKMTVVKLISHRECQOPHYYSGE 334
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
380 VRTVCLPPADLQLPDWTECSELSYGKHEALSPYSEKLEAHVRLYPSSRCTSQHLLNRT 439

QY 335 VITYKMLCAAD-----PQWKT-DSQCDSGGPLVCSIQGRMTLTGIVSWGRCALKDKPGV 388
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
440 VTDNMLCAGDTRSGGPGQANLHDACQDGGPLVCLNDGRMTLVGLIISWGLGCGQKQVPGV 499

QY 389 YTRVSHFLPWIRSHTK 404
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 YTKVTNYLDWIRDNR 515

RESULT 5
Q8C6L2 PRELIMINARY; PRT; 231 AA.
ID Q8C6L2;
AC Q8C6L2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator (Fragment).
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RX MEDLINE=22354683; Pubmed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK054349; BAC35743.1; -.
FT NON_TER 231
SQ SEQUENCE 231 AA; 25510 MW; 25E8980A682737F2 CRC64;

```

```
Query Match      36.6%; Score 842.5; DB 11; Length 231;
Best Local Similarity 71.5%; Pred. No. 4.7e-73;
Matches 143; Conservative 23; Mismatches 33; Indels 1; Gaps 1;

QY 9 SNCDCLNGGTCVSNKYFSNIHWCNCPKKGQHQHCEIDKSKTCYEGNGHFGKASTDTMG 68
Db 30 SNCGCQNGGVCSYKVFSTRRCSPKXFOGHECEIDASKTCYHGNGDSYRGKANTDTKG 89

QY 69 RPCLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQGLPLVQECMV 128
Db 90 RPLAWNAPVLOKPNYNAHRPDALSLGLGKHNYCRNPDNQKRPWCYVQIGLRFQVQECMV 149

QY 129 HDCADGKPPPEELKFCQGGQKTLRPFKIIGSEFTTIENQWFAAIYRRHGGSS_VTY 187
Db 150 HDCSLSKPSSVDQGGQCGQKALRPFKIVGSEFTTEVENQWFAAIYKKNKGGSPPSF 209

QY 188 VCGGSLISPCWISATHCFI 207
Db 210 KCGGSLISPCWASAAHCFI 229

RESULT 6
Q91VP2 PRELIMINARY; PRT; 559 AA.
AC Q91VP2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSSP; P00761; 1ANI.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD00018; KRINGLE.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; EGF; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRINCTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS02400; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
SQ Serine protease.
SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match      35.9%; Score 826.5; DB 11; Length 559;
Best Local Similarity 35.7%; Pred. No. 5.1e-71;
Matches 176; Conservative 65; Mismatches 155; Indels 97; Gaps 10;

QY 3 ELHVP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKKGQHQHCEIDKSKTCYEGNGHY 58
Db 74 QCHSVPRSCSEPRCNGGTCQALIFSDP-VQCQPDGFVGRKCDIDTRATCFEEQGIT 132

QY 59 RGKASTDTMGRPCLPWNASATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRNRRPWCYVQVG 118
Db 133 RGTWTAESGAECINWNSVLSUKPYNARRPNAIKLGLGNHNYCRNPDRLKPCWYVFKA 192

QY 119 LKPLVQECMVHDCADGKPS----- 138
Db 193 GKYTFECSTPACPKGKSEDCYVKGVTYRGTHSLTTSQASCLPWNISVLMGKSYTAWRT 252

QY 139 -----SPPEELK-----PQCG-OKTLRPRFKIIGG 162
Db 253 NSQALGLGRHNYCRNPDGDARPWCHVMKDKLTWECYDCMSPCSTGLRQYKRFQFRKGG 312

QY 163 EFTTIENQWFAAIY-RRHRGGSVTVYVCGSLISPCWISATHCFIDYPKKEDYIVYLGR 221
Db 313 LYTDITSHPWQAPIFVKNKRSPGERELCGVLISSCWVLSAAHCFLEPRFPNHLKVVLGR 372

QY 222 SRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALLLKIRSKEGRCAPSRITQICLP 281
Db 373 TVRVVPEBEQTEIEKYIVHEFDDT--YNDIALQLRSQSKQCAQESSVGTACL 430

QY 282 SMYNDPQF----GTSCEITGFGKENSTDYLYPQLKMTVVKLISHRECCQPHYVGEVTT 337
Db 431 ----DNLQLPDWTCELSGKGHEASSPFFSDRLKEAHVRLYPSSRCTSOHLFNKVTN 486

QY 338 KMLCAADP-----QWXTDCQDGGSLPVCSLQGRMTLTGIVSWGRGCAIKKPGYVTRV 392
Db 487 NMLCAGDTRSGNQDLHDACQDGGSLPVCMLINKQMTLTGIIISWGLGCGKQDVPGVYTKV 546

QY 393 SHFLPWIRSHTK 405
Db 547 TNYLDWIHDMKQ 559

RESULT 7
Q8MKB1 PRELIMINARY; PRT; 564 AA.
AC Q8MKB1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M, Yoshida E, Anai K, Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
```

```
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6DA937C CRC64;

Query Match 35.6%; Score 819; DB 6; Length 564;
Best Local Similarity 35.7%; Pred. No. 2.8e-70;
Matches 178; Conservative 55; Mismatches 148; Indels 118; Gaps 9;

QY 5 HOVP-SNCD---CLNGGTCVSNKYFNIHWCNCPKFGGQHCEID-----45
DB 80 HSPVSCSEPRCLNGTCSQALYFSDP-VCCPEGFVGRCEVDTRACYEDRGIGYRG 138
QY 46 -----45
DB 139 TWSSTESGAQCVMNWSLALPKYPYSGRKENALRLGLGNHNYCRNPDRDTKPNCYVFRAGT 198
QY 46 -----KKTCEYGNHGYRGKASTDTMGRLPCLPWSATVLTQYTHAHRSDA 91
DB 199 YSPFCSTPACSKENKNGCYLGGQAYRGTHSLTTSGLSCPLPWSMLLVGEKYTEARQNSA 258
QY 92 LQGLGKHNYCRNPONRRPWCVVQVGLKPLVQECMVHDCADGKXPSPPPEELKFCQG-Q 150
DB 259 EALGLGKHNYCRNPDGSDKWPCHVKNRLVIEYCDVPCA-----TCGR 304
QY 151 KTLRPRFKIIGEFTTIENQPFPAIY-RRHGGSVTYVCGSLISPCWVISATHCFIDY 209
DB 305 QDKQPFRIKGLGFTDITAHWPQAAIFVNNRSPGRFLCGGILINSCWVLSAAHCFLE 364
QY 210 PKEDYIVVYVGRSLNSNTQGMKEVENLILHKDYSADTLAHNDIALLKIRSKRGCA 269
DB 365 FPOQKRLVILGRTYPLVSAEEQIPEVEQPIIHERFDEGT--YNDIALLLKSTSGSCA 422
QY 270 QSRITQITCLPSMYNDPQGTSCITGFGKENSTDYLPQOLKMTVVKLISHRCQQPH 329
DB 423 QESQAVRLVCLPDASLQLPDWTCELSGKGKHEESPFSBQLKAHVLYPSSRCTPQQ 482
QY 330 YGSEVTTKMLCAADPQW-----KTDSQCGSGGLVCSLQGRMTLTGIVSWGRCALK 383
DB 483 LKNRTVTGNMLCAGDTRSGGAQWNLHDAQQGDSGGLVCMTDGHTLIGIISWGLGCGOK 542
QY 384 DKPGVYTRVSHPLPWIRSH 402
DB 543 DVPGVYTKVNYLGIQHH 561

RESULT 8
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
```

"T-plasminogen activator in tooth tissues."; Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases. -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.

RT RL
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR HMBL; AF364605; AAM00297.1; -.
DR HSSP; P00761; IAN1.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; FibTncnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR ProDom; PD000395; kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9B6B4C7CB101E8 CRC64;

Query Match 35.5%; Score 816.5; DB 6; Length 562;
Best Local Similarity 35.2%; Pred. No. 4.8e-70;
Matches 172; Conservative 61; Mismatches 166; Indels 89; Gaps 8;

QY 3 ELHQP-SNCD---CLNGGTCVSNKYFNIHWCNCPKFGGQHCEIDKSKTCYEGNHFY 58
DB 77 QCHSVFVSCSPRCFNGGTCLQALYFSDP-VCCPVPVFGIGRCEIDARATCYEDQGIT 135
QY 59 RKASCTDTMGRLPWSATVLTQYTHAHRSDALQGLGKHNYCRNPDRRRPWCVVQVG 118
DB 136 RGTWSTTESGAFCVNVNTSGLASMPYNGRPPDAVLGLGNHNYCRNPDKDSPWCYIFKA 195
QY 119 LKPLVQECMVHDCADGKXPSPPPEELKFO-----147
DB 196 EKSPDPFCSTPACTKEKEBECYTGKLDYRGTRSLTMSGAFCLPWISLVLGKIYTAWSN 255
QY 148 -----CGQKTLR-PRFKIIGE 163
DB 256 ACTLGLGKHNYCRNPDGDTQPWCHVLDKHKLWTEYCDLPQCVCGLRQYKEPQFRKGG 315
QY 164 FTTIENQPFPAIY-RRHGGSVTYVCGSLISPCWVISATHCFIDYPKKDYIVVYVGRS 222
DB 316 YADITSHWPQAAIFVNNRSPGRFLCGGILINSCWVLSAAHCFQERPPHVRVVLGRT 375
QY 223 RLNSNTQGMKEVENLILHKDYSADTLAHNDIALLKIRSKRGCAQPSRTIQTICLPS 282
DB 376 YRLVPEEEQAFVEKEYIVHKEFDDDT--YNDIALQLKSDSLTCAQESDAVRTVCLPE 433
QY 283 MYNDPQGTSCITGFGKENSTDYLPQOLKMTVVKLISHRCQQPHYVGSSEVTTKMLCA 342
DB 434 ANLQLPDWTECELSGKGKHEASSPFSYERLKAHVRLYPSSRCTSKHLFNKTIITNNMLCA 493
QY 343 ADPQW-----KTDSQCGSGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
DB 494 GDRSGGDNANLHDAQQGDSGGLVCMKGNHMTLVGVISWGLGCGOKDVPGVYTKVNYL 553
QY 397 PWIRSHTK 404
DB 554 NWIRDNTR 561

RESULT 9


```
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR00742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase;
KW Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 31.6%; Score 726; DB 11; Length 653;
Best Local Similarity 37.3%; Pred. No. 3.2e-61;
Matches 158; Conservative 57; Mismatches 167; Indels 42; Gaps 9;

OY 5 HQVPSNCDLNGGTCVSNKYFSNIHW-----CWCPRKFGQHCIEDKSKTCYEGNGH 56
DB 239 HTACLSSPCLNGGTC-----HLVGTGTSVCTCPYAGRCNIVPTHEHFLNGT 289
OY 57 FYRGKASTDTWGRPCLPNWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRERREWCVYQ 116
DB 290 EYRGVASTAAGSLSCLANWSLLYQELHVDVSAVAALLGLGPHAYCRNPDKDERPWCYV 349
OY 117 VGLKPLVCEQVHDCAD-GKRPSSPEEL-----KFCQGGK-----TLRPRFII 160
DB 350 KDNALSWCYRLTACESLARVHSQSPFILAALPESAPAVRPTCGRRHKRTFLRPR--II 407
OY 161 GGEFTTIENQPMFAAIYRRHRGGSVTVYCGSLISPCWVISATHCFIDVPKEDIVVLG 220
DB 408 GGSLSLPGSHPLAIIY---IGNS---FCAGSLVHTCWVSAHAFNAPSFRDSITVVLG 461
OY 221 RSLNSNTQGMKFEVNLILHKDYSADTLAHNDIALLKIRSEGRCAQPSRTITQICL 280
DB 462 QHFNRTTDTQTGIEKYVPVTLXSVFNENN-HLVLRLKKKGERCAVRSQFQPICL 520
OY 281 PSMYNDPQFGSCBTGFGKENSIDYLPQELKMTVVKLISHRECCQPHYGSEVTTKML 340
DB 521 PEAGSSFTGHKQCQIAGHGHMDENVSSYSNLSLEALVPLVADHKCSSPEVYGADISPNWL 580

OY 341 CAADPQWKTDSCQSDSGGLVCSLQRMWLTGIVSWGRGALKDKFGVTVRVSHFLPWIR 400
DB 581 CAGYFDCKSDACQSDSGGLVCEKNGVAYLYGIISWGGCGRLNKPGVTVRVANYVDWIN 640
OY 401 SHTK 404
DB 641 DRIR 644

RESULT 12
O97507 PRELIMINARY; PRT; 616 AA.
AC O97507;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Takahashi T., Kihara T.;
RT "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR00742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0013; FNTYPEII.
DR PRINTS; PRO0018; KRINGLE.
DR ProDom; PD000395; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00023; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBD1A9 CRC64;

Query Match 30.1%; Score 693.5; DB 6; Length 616;
Best Local Similarity 35.3%; Pred. No. 4e-58;
Matches 161; Conservative 61; Mismatches 159; Indels 75; Gaps 14;

OY 6 QVPSNCDLNGGTCVSNKYFSNIHWNCPCPKFGQHCIEDKSKTCYEGNGHFRGKASTD 65
```

```
Db 176 QVCTNPCLNGSGCLOTE---GHRLCRCPTGYAGRLCDVLDKERCYSDRGLSYRGAQTT 232
QY 66 TMRGPCLPMSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLKPLVQ 124
Db 233 LSGAPCPWAS----EATYWNMTAEQALNWLGDHAFCRPNDDTRWCFWFGDQLSWQ 288
QY 125 ECMVHDCAD-GKKP-----SSPPEEL-144
Db 289 YCLARCAPIGEAPILLTPTQSPSEHQDSPILLSREPOPTTQPSQNLTSAMCAPPBORG 348
QY 145 -----KFOCGQKTLRPF---KIIGBEFTTIENQWFAAIYRRHRGGSVTVCGSGLI 194
Db 349 PLPSAGLVGCGQR-LRRLSSLNRIVGLVALPOAHPIYIAALYWGQN-----FCAGSLI 401
QY 195 SPCWVISATCFIDYPKEDYIVYGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNN 254
Db 402 APCWLVTAACHLQNRPAPEELTVVLGQDRHNQSCQCTLAVERSYRLHESYSPKTYQH-- 459
QY 255 DIALLKTR-SKEGRCAQPSRTIQTICLP---SMYNDPQFGTSCBITGFKENSTDYLYPE 310
Db 460 DLALVRLKETADGCAHPSFPVQVCLPRSVASSAEPE-GALCEVAGWHQFEGAEYSS 518
QY 311 QLKMTVVKLISHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSGGPLVC---SLQGR 367
Db 519 FLQEAQVPLISPRCSAADVHGAAFTPGMLCAGFLEGGTDACQDSGGPLVCEDETAERQ 578
QY 368 MTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSH 403
Db 579 LVLRGIVSWGCGDRCLKPGVYTDVANYLAWIQEHT 614

RESULT 13
Q81Z55 PRELIMINARY; PRT; 615 AA.
AC Q81Z55;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR AB095845; BAC23095.1; -
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 29.9%; Score 689; DB 4; Length 615;
Best Local Similarity 34.3%; Pred. No. 1.1e-57;
Matches 153; Conservative 59; Mismatches 164; Indels 70; Gaps 9;

QY 13 CLNGGTCVSNKYFNTHWNCNPKKFGQHCIEDKSKTCYEGNGHFRGKASTDMGRPCL 72
Db 183 CLHGGRCLE---VEGHRLCPCPVGYTGFPCDVTDKASCYDGRGLSVRLGRTLLSGAPCQ 239
QY 73 PMSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLKPLVQECMVHDC 131
Db 240 PWAS----EATYENVTAEQARNWGLGHAFCRPNDDIRPWCFLNDRLSWEYCDLAQC 295
QY 132 -----ADGKFPSSPEEL 144
Db 296 QTPTQAPPTFVSPRLHVLMPAQAPPKPQPTTRTPQSQTPGALPAKREQPSLTRNG 355
QY 145 KFOGQ---KTLRPFKLIIGBEFTTIENQWFAAIYRRHRGGSVTVCGSGLISPCWVIS 201
Db 356 PLSCGRLRKSLSMTKRVVGLVALRGHAHPYIAALYWGHS-----FCAGSLIAPCWLT 409
QY 202 ATHCFIDYPKEDYIVYGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKI 261
```

```
Db 410 AAHCLQDRPAPEDLTIVLQGERRNHSCPCQTLAVRSYRLHEAFS---PVSQYHDLALLRL 467
QY 262 R-SKEGRCAQPSRTIQTICLPMSMYNDPQFGTSCBITGFKENSTDYLYPEOLKMTVVVKLI 320
Db 468 QEDADSCALLSPYQVPCLPSPAARPSETTLQVACGQHGFEGAEYASFLQEAQVFFL 527
QY 321 SHRECOQPHYVGSVTTKMLCAADPQWKTDSCQDSGGPLVCSLQG---RMTLTGIVSWG 377
Db 528 SLERCAPDVHGSILPGMLCAGFLEGGTDACQDSGGPLVCEQAAERLTLQGIISWG 587
QY 378 RGALXDKDPQVYTRVSHFLPWIRSH 403
Db 588 SSGCDNRNKPVGYYTDVAYLAWIREHT 613

RESULT 14
Q9TVAB PRELIMINARY; PRT; 157 AA.
AC Q9TVAB;
DT 01-MAY-2000 (T-EMBLrel. 13, Created)
DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA Balcerzak D., Querengesser L., Dixon W.T., Baracos V.E.;
RT "Coordinate expression of matrix-degrading proteinases and their
activators and inhibitors in bovine skeletal muscle.";
RL J. Anim. Sci. 79:94-107(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF1474761; AAD30301.1; -.
DR HSP; P00749; IURK.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PSS0070; KRINGLE_2; 1.
KW Glycoprotein; Kinase; Kringle.
FT NON_TER 1
FT NON_TER 157
SQ SEQUENCE 157 AA; 17858 MW; A768D6C72C1FBFB7 CRC64;

Query Match 29.4%; Score 677; DB 6; Length 157;
Best Local Similarity 73.9%; Pred. No. 2.7e-57;
Matches 116; Conservative 15; Mismatches 26; Indels 0; Gaps 0;

QY 13 CLNGGTCVSNKYFNTHWNCNPKKFGQHCIEDKSKTCYEGNGHFRGKASTDMGRPCL 72
Db 1 CLNGGKCVTKYKPSNTRQSCPKKFGQHCIEDTSKTCYEGNGHFRGKANDLSGRPCL 60
QY 73 PMSATVLOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLKPLVQECMVHDC 132
Db 61 AMDSPVLLKMGVHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLKPLVQECMVDCS 120
QY 133 DGKFPSSPEELKFCQCGKTLRPFKLIIGBEFTTIEN 169
Db 121 VGKSPSSPREKBEFQCGQKALRPFKIVGGQVTNAEN 157

RESULT 15
O35727
```

ID	Q35727	PRELIMINARY;	PRT;	597 AA.
AC	Q35727;			
DT	01-JAN-1998	(TrEMBLrel. 05, Created)		
DT	01-JAN-1998	(TrEMBLrel. 05, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	Factor XII.			
GN	P12.			
GN	P12.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=liver;			
RA	Schloesser M., Schwager S., Engel W.;			
RL	Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; X99571; CAA67891.1; -.			
DR	HSSP; P00760; 1A07.			
DR	MEROPS; S01.211; -.			
DR	MGI; 1891012; F12.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR006209; EGF like.			
DR	InterPro; IPR000083; Fibnctn1.			
DR	InterPro; IPR000562; FN Type II.			
DR	InterPro; IPR006210; IEGF.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00008; EGF; 2.			
DR	Pfam; PF00039; fn1; 1.			
DR	Pfam; PF00040; fn2; 1.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00013; FNTYPEII.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000995; FN Type II; 1.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00181; EGF; 2.			
DR	SMART; SM00058; FN1; 1.			
DR	SMART; SM00059; FN2; 1.			
DR	SMART; SM00130; KR; 1.			
DR	SMART; SM00020; Tryp_spc; 1.			
DR	PROSITE; PS00022; EGF_1; 2.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.			
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS00070; KRINGLE_2; 1.			
DR	PROSITE; PS02440; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;			
KW	Serine protease.			
SEQUENCE	597 AA; F3AC07C37D0C0FBA CRC64;			
Query Match	29.4%; Score 677; DB 11; Length 597;			
Best Local Similarity	34.9%; Pred. No. 1.5e-56;			
Matches 149; Conservative	64; Mismatches 162; Indels 52; Gaps 10;			
QY	13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72			
Db	183 CLNGGSL---LVEDHPLCRPTGYGYFCDDLWATCYEGRLSVRGAGTQSGAPCQ 239			
QY	73 PMSATVLOQTY-HAHRSDALQGLGKHNYCRPNDRPCVYQVGLKPLVQECMVHDC 131			
Db	240 RW----TVEATYRNMTKEQALSWGLGHAFCRPNDRPNDTRPCVFWSGDRLSDYCGLEQC 295			
QY	132 -----ADGKKPSPP-----EELKFCQCG---KTLRPRFKI 159			
Db	296 QTFTFAPLVVPSQESQAPSLSHAPNDSDTHQTSLSKNTMTMGGQGRFKGLSFMKV 355			
QY	160 IGGEFTTIENQWPFAAIYRRHGRGSVTYVCGGSLISPCWVSWATHCFIDYVKKEDIYVL 219			
Db	356 VGLGVALPGSHPIYAAIYWGNN-----FCAGSLIAPCWLTAAHCLQNRPAPEELTVVL 409			
QY	220 GRSRLNSNTQGMKFEVENLLHKDYSDFTLAHNDIALKIR-SKEGRCAQPSRTIQT 278			
Db	410 GQDRHNQSCWECOTLAVRSYRLHEGFSITYQH--DLALLRLQESKTNSCAILSPHVQPV 467			
QY	279 CLFSMYNDPQFGTSCITGFGKENSTDYLYPEOLKMTVVKLISHRECCQPHYGVSEVTTK 338			
Db	468 CLFSGAAPPSETVLCVAGWGHLQEGAEFESTLQEAQVFFIALDRCSNSNVHGDALPG 527			
QY	339 MLCADDPQWKTDSCQDSDSGPLVC---SLQGRMTLTGIVSWGRGCALKDKPGVYRVSHF 395			
Db	528 MLCAGFLEGGTDACQDSDSGPLVCBEGTABHQTLRGVLSWGGCGDRNKPVGYYTVVANY 587			
QY	396 LPWIRSH 402			
Db	588 LAWIQKH 594			
RESULT	16			
Q8XOD2	PRELIMINARY;	PRT;	517 AA.	
ID	Q8XOD2			
AC	Q8XOD2;			
DT	01-OCT-2002	(TrEMBLrel. 22, Created)		
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)		
DE	Hypothetical protein.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RA	Strausberg R.;			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.			
DR	EMBL; BC031775; AAH31775.1; -.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR000742; EGF_2.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	InterPro; IPR006209; EGF like.			
DR	InterPro; IPR006210; IEGF.			
DR	InterPro; IPR000001; Kringle.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00008; EGF; 2.			
DR	Pfam; PF00051; kringle; 1.			
DR	Pfam; PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	PRINTS; PR00018; KRINGLE.			
DR	ProDom; PD000395; Kringle; 1.			
DR	SMART; SM00181; EGF; 2.			
DR	SMART; SM00058; FN1; 1.			
DR	SMART; SM00059; FN2; 1.			
DR	SMART; SM00130; KR; 1.			
DR	SMART; SM00020; Tryp_spc; 1.			
DR	PROSITE; PS00022; EGF_1; 2.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.			
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.			
DR	PROSITE; PS00021; KRINGLE_1; 1.			
DR	PROSITE; PS00070; KRINGLE_2; 1.			
DR	PROSITE; PS02440; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;			
KW	Kringle; Protease; Serine protease.			
SEQUENCE	517 AA; 3855A42035A5EA59 CRC64;			
Query Match	29.3%; Score 674.5; DB 11; Length 517;			
Best Local Similarity	38.0%; Pred. No. 2.2e-56;			
Matches 159; Conservative	58; Mismatches 164; Indels 37; Gaps 16;			
QY	13 CLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72			

```
Db 116 CQNGGVCRRHRRSRF--TCACPDQYKGFCEIGPD-DCYVGDGYSYRGKSKTVNQNPCL 173
Qy 73 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYQVGLKPLVQE--CMVHDC 131
Db 174 YNSHLLQLQTYNFMEDAEATHGIAEHFNCENPDGDHKPWCYKVNSEKVKWEYCDVTVC 233
Qy 132 ADGKKPSPPEEL-----KFG-CQOKTLRPRF--KIIGETTTIENQFWFAAIY--- 177
Db 234 PVPDTP-NPVESLLEPVMELPFFSCGKTEVAEHAHVRIYGGFKSTAGKHPQWVSLQSL 292
Qy 178 ---RRHGGSVTVVCGSLISPCWVISAHCFIDYPKEDYIVYLGSRSLNSNTQGEK 233
Db 293 PLTTSMFQG---HFCCGALHPCWVLTAAHC-TDINTKHLKV-LGDQDLKKTESHEQT 346
Qy 234 FEVENLILHKDYSADTLAHNDIALLKIRSEKGRCAQPSRTIOTICLPSMYNDP-QFGTS 292
Db 347 FRVEKILKYSQYNERDEIPHNDIALLKLPVGGHCALESRYVKTVCLPS---DFPSTGTE 403
Qy 293 CEITGFGKENSTLYPEOLKMTVVKLISHRECQOPHYGSEVTTMLCAADPQWK-TDS 351
Db 404 CHISGMGVTETGE--GSRQLLDAKVLIANPLCNSRLYDHTIDDSMICAGNLQKPGSDT 461
Qy 352 CQSDGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHELPIRSHTKKEENG 409
Db 462 CQSDGGGLTCEKDGTYVYVYVGVSWGQECG--KFGVYTVQTKFLNWKTMHREAGL 517

RESULT 17
Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA Kitamura N.;
RP Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Miura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma; it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator.";
RL J. Biochem. 119:1157-1165 (1996).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
RP Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -
DR EMBL; S83182; AAB46909.1; -
DR EMBL; BC031412; AAH31412.1; -
DR HSSP; P00763; IDPO.
DR MEROPS; S01.033; -
DR Genew; HGNC:4798; HAP2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_Protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
```

```
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR PROSITE; PS00022; EGF_1; 3.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 28.7%; Score 661.5; DB 4; Length 560;
Best Local Similarity 37.1%; Pred No. 4.4e-55;
Matches 155; Conservative 58; Mismatches 166; Indels 39; Gaps 15;

Qy 13 CLNGGTCVSNKYFSNIHWCNCPKFGQGHCEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 159 CQNGATCSRHKRSKF--TCACPDQYKGFCEIG--SDDCYVGDGYSYRGKMNRTVNOHACL 216
Qy 73 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYQVGLKPL-----VQEC 126
Db 217 YNSHLLQLQTYNFMEDAEATHGIEHNFNCNPDADKPCWCFIKVTNDKVKWEYCDVSAC 276
Qy 127 MVHDCA-DGKKPSPPEELK--FQCGOKTLRPR--FKIIIGEFTEIENQFWFAAIYRRHR 181
Db 277 SAQDVAYPEESPTEPSTKLPFGDSCGKTEAERKIKIYGGFKSTAGKHPQASLQ--- 332
Qy 182 GGSVT-----YVCGSLISPCWVISAHCFIDYPKEDYIVYLGSRSLNSNTQGEK 233
Db 333 -SSLPLTISMPQGHFCGALIHPCWVLTAAHC-TDI-KTRHLKVVLGDQDLKKEEFHEQS 389
Qy 234 FEVENLILHKDYSADTLAHNDIALLKIRSEKGRCAQPSRTIOTICLPSMYNDPQF--GT 291
Db 390 FRVEKIFKYSHYNERDEIPHNDIALLKLPVGGHCALESRYVKTVCLP---DGSFPSSG 445
Qy 292 SCEITGFGKENSTLYPEOLKMTVVKLISHRECQOPHYGSEVTTMLCAADPQWK-TD 350
Db 446 ECHISGMGVTETGK--GSRQLLDAKVLIANPLCNSRLYDHTIDDSMICAGNLQKPGQD 503
Qy 351 CQSDGGGLVCSLQGRMTLTGIVSWGRCALKDKPGVYTVRSHELPIRSHTKKEENG 408
Db 504 TCQSDGGGLTCEKDGTYVYVYVGVSWGLEG--KFGVYTVQTKFLNWKATIKSESG 559

RESULT 18
O97587 PRELIMINARY; PRT; 128 AA.
AC O97587;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand White;
RX MEDLINE=99057575; PubMed=9837780;
RA Reno C., Boykiw R., Martinez M.L., Hart D.A.;
RT "Temporal alterations in mRNA levels for proteinases and inhibitors
RT and their potential regulators in the healing medial collateral
RT ligament.";
RL Biochem. Biophys. Res. Commun. 252:757-763 (1998).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AF069711; AAC95003.1; -
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; -.
```

```

DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR001254; Ser. protease_Try.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SMO0020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1 128
FT NON_TER 128 128
SQ SEQUENCE 128 AA; 14328 MW; 1BC7ED30E071A06D CRC64;

Query Match 25.4%; Score 585; DB 6; Length 128;
Best Local Similarity 84.4%; Pred. No. 1.6e-48;
Matches 108; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 172 WFAAIYRRHGGSVTVYCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNTQGE 231
Db 1 WFAAIYRRHGGSVTVYCGSLSPCWVISATHCFINHQKEDYIVYLGSRSLNSNTPGE 60

Qy 232 MKFEVENLILHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGT 291
Db 61 MKFEVEQLLHGYRADTLAHNDIALKILSNNGCAQPSRSIQTICLPPWNADPNFGT 120

Qy 292 SCEITGFG 299
Db 121 SCEITGFG 128

RESULT 19
Q91WJ5 PRELIMINARY; PRT; 812 AA.
AC Q91WJ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen.
GN PLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,
RA Nagaraja R.;
RT "Genomic Sequence Analysis in the Mouse t-complex Region.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; BC014773; AAH14773.1; -.
DR EMBL; AF481053; AAM22156.1; -.
DR HSSP; P00761; 1ANI.
DR MGD; MGL:97620; Plg.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR InterPro; IPR001400; Somatostatin.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR0018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR Predom; PD000395; Kringle; 5.
DR PROSITE; PS00021; KRINGLE_1; 5.

DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS00338; SOMATOSTATIN_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 812 AA; 90781 MW; 2417326086A2FFD2 CRC64;

Query Match 21.8%; Score 501; DB 11; Length 812;
Best Local Similarity 32.5%; Pred. No. 2.2e-39;
Matches 135; Conservative 48; Mismatches 170; Indels 62; Gaps 15;

Qy 10 NCDCLNGGTCTVSNKYPSNIHWCNCPK--KFGQHQCEI-----DKSKTCVEGNHG 56
Db 428 NPDGDKGWCYTTDPSRWWEYCNLXRCSETGGSVVLELPTVSQSPGSDSETDCMYGNKG 487

Qy 57 FYRGKASDTMGRPCLPWNSATVLQOTVHAHRSALQ---LGLKHNKYNPD-NRRRP 111
Db 488 DYRGKATVTAAGTTCQGWAA-----QEPHRSIFTPQTNPRAGLEK-NYCRNPDGVNPG 541

Qy 112 WCYVVGVLKPLVQECMVHDCADGKKPSPBPBELKFCQCKTLRPR---FKIIGSBFTTIE 168
Db 542 WCYT-TNPKLYDYCDIPLCASAS-----SPECGKQPVEPKKCPGRVVGVCVANPH 591

Qy 169 NOPWFAAIYRRHGGSVTVYCGSLSPCWVISATHCFIDYPKKEDYIVYLGSRSLNSNT 228
Db 592 SWPQWISLRTFTG---QHFCGGTLLIAPEWLVTAHCKLEKSSRPFEYKVLGAHEEYIRG 648

Qy 229 QGEMKFEVENLILHKDYSADTLAHNDIALKIRSEGRCAQPSRTIQTICLPSMYNDPQ 288
Db 649 SDVQEISVAKLLE-----PNNRDIALKL---SEPATITDKVIPACLFPS---PN 692

Qy 289 P----GTSCEITGPKENSTLYPEQLKNTVVKLIHRECOQPHYSEVTTKMLCAAD 344
Db 693 YMVADRTICITGWGETGT--FGAGRLKBAQLVIENKVCNRYEVLNNRVAKSTELCAGQ 750

Qy 345 PQWKTDSCQDGGGLVCSLQGRWTLTGIVSVGRGKALKDKPGVTVTRYSHFLPWI 399
Db 751 LAGVDSCQDGGGLVCSLQGRWTLTGIVSVGRGKALKDKPGVTVTRYSHFLPWI 805

RESULT 20
Q15146 PRELIMINARY; PRT; 810 AA.
AC Q15146;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Browne M.J., Chapman C.G., Dodd I., Carey J.E., Lawrence G.M.P.,
RA Mitchell D., Robinson J.H.;
RT "Expression of recombinant human plasminogen and aglycoplasminogen in
RT HeLa cells.";
RL Fibrinolysis 0:0-0(1991).
CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; M74220; AAA36451.1; -.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan-app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.

```

```
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR PR0DOM; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRYD SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS50070; KRINGLE 2; 5.
DR PROSITE; PS02040; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 810 PLASMINOGEN.
SQ SEQUENCE 810 AA; 90555 MW; B05C7D4B0D020B3C CRC64;

Query Match 21.6%; Score 498; DB 4; Length 810;
Best Local Similarity 33.3%; Pred. No. 4.2e-39;
Matches 138; Conservative 45; Mismatches 168; Indels 64; Gaps 16;

QY 10 NCDCNLGGTCVSNKYFNSIHKNCPKKFGGQ-----HCEIDKSKTCYEGNGH 56
Db NPDAKGPWCFTTDPVSRWBYCNLKKCSGTEASVAPPVVLVLPDVTPEEDCMFGNGK 487
QY 57 FYRGKASTDTMGRPCLPWNSATVLQOYYHAHR-----SDALQLGLGKHNYCRNPD-NRRRP 111
Db GYRGKRAITVTFPCQDAA-----QEPHRHSITFTETNPRAGLEK-NYCRNPDGDVGGP 541
QY 112 WCYVQVGLKPLVQRCMVHDCADGKKSPSPPEELKFCQCKTLRPR---FKIIGGEFTTIE 168
Db WCYT-TNPRKLYDCYDFQCA---APS-----FDCGKQVPEKPCGPRVGGCVVAHPH 590
QY 169 NQPFAAIYRHRGGSVTYVCGGSLISPCWVISATHCFIDYPKKEDIYVILGRSLNSNT 228
Db SNFWQSL--RTRPG--MHFCGGTLLSPWVLTAAHCLKSPRPSSYYKVLGAHQEVNLE 646
QY 229 QGEMKFEVENLILHKDYSADTLAHHNDIALKIRSEKGRCAQPSRTTQICLPSMYNDPQ 288
Db PHQEIERSVLFLEPT-----RKDIALLKLSPP---AVITDKVIPACLPFS----FN 690
QY 289 F----GTSCEITGFKENSNDYLYPEOLKMTVVVKLISHRECQOQPHYGVSEVTTKMLCAAD 344
Db YVVADETRCEFTVGTGETQGT--FGAGLLKEAQLPVIENKVCNRYEFLNGRVQSTELCAGH 748
QY 345 POKWTDSCQDGGPLVCSLQGRMTLGIYVSWRGCAKDKPKGYTVVSHFLPWI 399
Db LAQGTDSQQDGGPLVCFKDKYILQGVTSWGLGCAKPNKPGYVYVRVSRFVTWI 803

RESULT 21
O46506 PRELIMINARY; PRT; 454 AA.
AC O46506;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein a (Fragment).
GN BABAPOA.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice Site Mutation is Associated with Deletion of a Single Exon in a Null Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
```

```
DR EMBL; AF029691; AAB97886.1; -.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.999; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PR0DOM; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; TRYD SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 2.
DR PROSITE; PS50070; KRINGLE 2; 2.
DR PROSITE; PS02040; TRYPSIN DOM; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Lipoprotein; Protease;
KW Serine protease.
FT NON_TER 1 1
SQ SEQUENCE 454 AA; 50041 MW; 974E30744C187B2F CRC64;

Query Match 21.6%; Score 496; DB 6; Length 454;
Best Local Similarity 34.0%; Pred. No. 3.1e-39;
Matches 130; Conservative 52; Mismatches 132; Indels 68; Gaps 15;

QY 50 CYEGNGHYFGKASTDTMGRPCLPWNSATVLQ--QTYHAHRSDALQLGLGKHNYCRNPDN 107
Db CYHGDGQSVRGFSFTVTGRTCSQSWSMTPHQHKRTPENHPNDGLTN----NYCRNPDA 156
QY 108 RRPWCYVQVGLKPLV--QECMVHDCAD-----GKKPSSPPEEL 144
Db DTGPWCFT---MDPSRWCEYCNLTRCSDTGTGVTPLVTIPIPSLEARSQASS----- 208
QY 145 KFCQGKTLRPR---FKIIGGEFTTIENQPFAAIYRHRGGSVTYVCGGSLISPCWVIS 201
Db FDCGKQVPEKPCGPRVGGCVVAHSAHWPQVSL--RTRFGK--HFCGGLISPEWVLT 263
QY 202 ATHCFIDYPKKEDIYVILGRSLNSNTQGMKEVEVENLILHKDYSADTLAHHNDIALK 261
Db AARCLEMSRPSSYYKVLG-----AHQVNLSEHVEIEVSKLFSEPTGA---DIALKL 315
QY 262 RSKEGCAQPSRTTQICLPSMYNDPQF---GTSCEITGFKENSNDYLYPEOLKMTVV 317
Db SPALITDKVIPACLPFS---PNYVITAMTECITGWTQGT--FGAGLLREARL 365
QY 318 KLISHRECQOQPHYGVSEVTTKMLCAADPQWKTDCQDGGPLVCSLQGRMTLTGVISWG 377
Db PVIENTVCNRYEFLNGRVKSTELCAGHLAGGTDSQQDGGPLVCFKDKYILRGITSWG 425
QY 378 RGCALKDKPGYTVVSHFLPWI 399
Db PGCAEPNKPQYVYVRVSRFVTWI 447

RESULT 22
Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
```


Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

```

RL  Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RC  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RA  MEDLINE=91250378; PubMed=1645711;
RA  Kanalas J.J., Makker S.P.;
RT  "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT  receptor site for plasminogen."
RL  J. Biol. Chem. 266:10825-10829(1991).
CC  -!- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR  EMBL; AJ242649; CAB46014.1; -.
DR  HSSP; P00747; 1PMK.
DR  InterPro; IPR001314; Chymotrypsin.
DR  InterPro; IPR000001; Kringle.
DR  InterPro; IPR003014; Pan.
DR  InterPro; IPR003609; Pan app.
DR  InterPro; IPR003966; Prothrombin.
DR  InterPro; IPR001254; Ser. protease_Try.
DR  InterPro; IPR001400; Somatostatin.
DR  Pfam; PF00051; Kringle; 5.
DR  Pfam; PF00024; PAN; 1.
DR  Pfam; PF00089; trypsin; 1.
DR  PRINTS; PR00722; CHYMOTRYPSIN.
DR  PRINTS; PR00018; KRINGLE.
DR  ProDom; PD000395; Kringle; 5.
DR  SMART; SM00130; KR; 4.
DR  SMART; SM00473; PAN AP; 1.
DR  SMART; SM00020; Tryp_SPC; 1.
DR  PROSITE; PS00021; KRINGLE 1; 5.
DR  PROSITE; PS00070; KRINGLE 2; 5.
DR  PROSITE; PS00338; SOMATOSTATIN 2; 1.
DR  PROSITE; PS00240; TRYPSIN_DOM; 1.
DR  PROSITE; PS00134; TRYPSIN_HIS; 1.
DR  PROSITE; PS00135; TRYPSIN_SER; 1.
KW  Glycoprotein; Hydrolase; Kringle; Protease; Serine protease; Signal.
FT  SIGNAL 19
FT  CHAIN 20 812 PLASMINOGEN.
SQ  SEQUENCE 812 AA; 90535 MW; 8C703C51410BEC9E CRC64;

Query Match 21.6%; Score 496; DB 11; Length 812;
Best Local Similarity 33.2%; Pred. No. 6.6e-39;
Matches 136; Conservative 56; Mismatches 148; Indels 70; Gaps 20;

QY 27 NIHW--CN---CPKFKG--QCEIDK-----SKT-CYEGNGHFYRGKASTDTMGPC 71
DB 27 NIHW--CN---CPKFKG--QCEIDK-----SKT-CYEGNGHFYRGKASTDTMGPC 71
QY 443 SVRWEYCNLKRCSGTGGGVAESAIVQVPSAPGTSETDCMTGNGKRGYRGKTAVTAAGTTC 502
DB 443 SVRWEYCNLKRCSGTGGGVAESAIVQVPSAPGTSETDCMTGNGKRGYRGKTAVTAAGTTC 502
QY 72 LPWNSATVLOQTYHAHRSALQ----LGLGKHNYCRNPD-NRRPWCYVQVGLKPLVQBC 126
DB 72 LPWNSATVLOQTYHAHRSALQ----LGLGKHNYCRNPD-NRRPWCYVQVGLKPLVQBC 126
QY 503 QEWAA-----QEPHSHRIFTPTNPRAGLEK-NYCRNPDGVNGPWCYT-MNPRKLYDYC 555
DB 503 QEWAA-----QEPHSHRIFTPTNPRAGLEK-NYCRNPDGVNGPWCYT-MNPRKLYDYC 555
QY 127 MVHDCADGKPSPPPEELKFOGOKTLRPR--FKLIGETTTIENQWFAAIYRHRGG 183
DB 127 MVHDCADGKPSPPPEELKFOGOKTLRPR--FKLIGETTTIENQWFAAIYRHRGG 183
QY 556 NIPLCASL-----SFCGRQVPEKCPKGVGVGCVANPWSHPWQISLRTFSG- 605
DB 556 NIPLCASL-----SFCGRQVPEKCPKGVGVGCVANPWSHPWQISLRTFSG- 605
QY 184 SVTVYCGSLSPCWVISAHCIFIDYPKKEDYIVYLGSR---LNSNTQGMKFEVENLI 240
DB 184 SVTVYCGSLSPCWVISAHCIFIDYPKKEDYIVYLGSR---LNSNTQGMKFEVENLI 240
QY 606 --QHFCGTLSPWVLTAAHCLSKSRPEFYKVLGAHEERILGSDVQ---QIATVTLV 660
DB 606 --QHFCGTLSPWVLTAAHCLSKSRPEFYKVLGAHEERILGSDVQ---QIATVTLV 660
QY 241 LHKQVSAATLAHNDIALLKIRSGRCQAQSPRTIOTCLPSMYNDPOF-----GTSCEIT 296
DB 241 LHKQVSAATLAHNDIALLKIRSGRCQAQSPRTIOTCLPSMYNDPOF-----GTSCEIT 296
QY 661 LEPNDA-----DIALKL-----SRPATITDNNVPACLPF-----ENYVAVDRLCVIT 704
DB 661 LEPNDA-----DIALKL-----SRPATITDNNVPACLPF-----ENYVAVDRLCVIT 704
QY 297 GFGENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTSCCGDS 356
DB 297 GFGENSTDYLYPEOLKMTVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTSCCGDS 356
QY 705 GWGETYKTP--GAGLKAQLPVIENKVNRAEYLNVRKSTELCAGHLAGIDSCQDS 762
DB 705 GWGETYKTP--GAGLKAQLPVIENKVNRAEYLNVRKSTELCAGHLAGIDSCQDS 762
QY 357 GGPLVCSLQGRWTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHYKKE 406
DB 357 GGPLVCSLQGRWTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHYKKE 406
QY 763 GGPLVCFEKDKVILQGVTSWGLGCAPKPKPGVYVRSYVNVNWEREND 812
DB 763 GGPLVCFEKDKVILQGVTSWGLGCAPKPKPGVYVRSYVNVNWEREND 812

```

```

RESULT 23
Q95EF3 PRELIMINARY; PRT; 300 AA.
AC Q95EF3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RS SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; BC012390; AAH12390.1; -.
DR HSSP; P00761; IANI.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; Hydrolase; Protease; Serine protease.
SQ SEQUENCE 300 AA; 32216 MW; FFC2BDF9382F636A CRC64;

Query Match 21.5%; Score 495.5; DB 4; Length 300;
Best Local Similarity 38.4%; Pred. No. 2.1e-39;
Matches 107; Conservative 45; Mismatches 112; Indels 15; Gaps 5;

QY 132 ADGKSPSPPEELKFOGQ---KTLRPFKIKGGFTTIENQWFAAIYRHRGGSVTV 188
DB 132 ADGKSPSPPEELKFOGQ---KTLRPFKIKGGFTTIENQWFAAIYRHRGGSVTV 188
QY 28 AKREQPSPLTRNGPLSCGQRLKSLSSMTRVVGLVALGAHPYIAALYGHG-----F 81
DB 28 AKREQPSPLTRNGPLSCGQRLKSLSSMTRVVGLVALGAHPYIAALYGHG-----F 81
QY 189 CGSLSPCWVISAHCIFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLIHKDYSAD 248
DB 189 CGSLSPCWVISAHCIFIDYPKKEDYIVYLGSRNSNTQGMKFEVENLIHKDYSAD 248
QY 82 CAGSLIAPCWVLTAAHCLQDRPAEDLTIVLQGERNHSCPCQTLAVRSYRLHEAFS-- 139
DB 82 CAGSLIAPCWVLTAAHCLQDRPAEDLTIVLQGERNHSCPCQTLAVRSYRLHEAFS-- 139
QY 249 TLAHNDIALLKIR-SKEGRCAQSPRTIOTCLPSMYNDPOQFGTSCETITFGKENSTDYL 307
DB 249 TLAHNDIALLKIR-SKEGRCAQSPRTIOTCLPSMYNDPOQFGTSCETITFGKENSTDYL 307
QY 140 PVSQDHLALLRLQEDADGSCALLSYQVCLPFGAARPSSETTLCQVAGWGHQEGAE 199
DB 140 PVSQDHLALLRLQEDADGSCALLSYQVCLPFGAARPSSETTLCQVAGWGHQEGAE 199
QY 308 YPEQLKMTVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTSCQDGGSGPLVCSLQ- 366
DB 308 YPEQLKMTVVKLISHRECQPHYGVSEVTTKMLCAADPQWKTSCQDGGSGPLVCSLQ- 366
QY 200 YASFLEAQVPEFLSLERCSAPDVHGSSILPGLMCAAGLEGGTDACQDGGSGPLVCEQAA 259
DB 200 YASFLEAQVPEFLSLERCSAPDVHGSSILPGLMCAAGLEGGTDACQDGGSGPLVCEQAA 259
QY 367 --RMTLTGVISWGRCALKDKPGVYTRVSHFLPWIRSHYKKE 403
DB 367 --RMTLTGVISWGRCALKDKPGVYTRVSHFLPWIRSHYKKE 403
QY 260 ERRLTQGIISWGSGCGDRKPKGVITDVAIYLAIREHT 298
DB 260 ERRLTQGIISWGSGCGDRKPKGVITDVAIYLAIREHT 298

```

```

RESULT 24
Q95M89 PRELIMINARY; PRT; 103 AA.
AC Q95M89;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen activator urokinase (Fragment).
GN PLAU.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RS SEQUENCE FROM N.A.
RL MEDLINE=21314992; PubMed=11421942;
RX Shubtowski D.M., Ventr P.J., Douglass C.L., Zhou R.-X., Ewart S.L.;
RA

```

RT "polymorphism identification within 50 equine gene-specific sequence
tagged sites.";

RL Anim. Genet. 32:78-78 (2001).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL: AY008806; AAK14840.1; .
DR EMBL: AY008803; AAK14840.1; JOINED.
DR InterPro: IPR001314; Chymotrypsin.
DR Pfam: PF00089; Trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS0134; TRYPSIN_HIS; 1.
KW Hydrolase; Kinase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 103
SQ SEQUENCE 103 AA; 11525 MW; 0B739514F6331180 CRC64;

Query Match 21.5%; Score 495; DB 6; Length 103;
Best Local Similarity 87.4%; Pred. No. 5.9e-40;
Matches 90; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
QY 172 WFAAIVRRHRGGVTVVCGSLISPCWVISATHCFIDYPKEDYIVYLGSRSLNSNTQGE 231
Db 1 WFAAIVRRHRGGVTVVCGSLISPCWVLSATHCFINYPKEDYIVYLGSRSLNSTSGE 60
QY 232 MKEVENLILHKDYSADTLAHNDIALLKIRSGRCQAQPSRT 274
Db 61 MKEVEKLILHEDYSADTLAHNDIALLKISSTGTGCAQPSRS 103

RESULT 25

Q8AVB0 PRELIMINARY; PRT; 429 AA.
AC Q8AVB0
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Plasminogen precursor (Fragment).
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Hanumanthaiah R., Day K., Jagadeeswaran P.;
RT "Comprehensive analysis of blood coagulation pathways in teleostei:
RT Evolution of coagulation factor genes and identification of zebrafish
RT Factor VIIi.";
RL Blood Cells Mol. Dis. 0:0-0(2002).
DR EMBL: AF515276; AAN71006.1; .
FT NON_TER 1
SQ SEQUENCE 429 AA; 47556 MW; 9A580A214A549C12 CRC64;

Query Match 21.4%; Score 491.5; DB 13; Length 429;
Best Local Similarity 35.3%; Pred. No. 8e-39;
Matches 130; Conservative 36; Mismatches 153; Indels 49; Gaps 14;
QY 48 KTCYEGNGHYRGKASDTDMGRCLPWN SATVLQOYHAAH---RSDALQLGLGKHNYCRNP 104
Db 100 KDCCKNGAIFYRGSTSMYGVTCQAWRSMTPHQASFTPTHPDKGLE-----SNQCRN 154
QY 105 PDNR-RRPWCYVQVGLKPLVQECWHDCA DGKPKSPPEELKFQCGQKTLRPR---FKII 160
Db 155 PDSVNGPWCYTTDPSKWK--DYCIQIDC-----ESLK--CQGPATPKRCGRIV 201
QY 161 GGEFTTIENQWPFAAIVRRHRGGVTVVCGSLISPCWVISATHCFIDYPKEDYIVYLG 220
Db 202 GGCVRKPHSPWP--QISLRTRG--KTHFCGGLTLDPOWVTAHCLERSDPSAYKIMLG 257
QY 221 RSLNSNTQGEKMEFVENLILHKDYSADTLAHNDIALLKIRSGRCQAQPSRTIOTICL 280

Db 258 IHTERTATESKQBRDVTKII---KGPA GT-----DIALKL-----DRPALINDKVPVCL 305
QY 281 PSMYNDPQFGTSCETITGFGKENST---DYLYPQLKMTVVVKLIASHRECOQPHYGVSEVTT 337
Db 306 PEKDIYVPSNTECYVTGWGETQDTGGEGY-----LKETGFPVIENKVCNRPFLNGRVKD 360
QY 338 KMLCAADPQWKTDSCQDSDGGLVCSLQGRMTLTGIVSGRGCALDKDKPGVYTRVSHPLP 397
Db 361 HEMCAGNIEGNDSQDSDGGLVCAQNTFVLQGVTSWGLGCANAMKPGVYTRVSKFVD 420
QY 398 WIRSHTKK 405
Db 421 WIRSKE 428

RESULT 26

O46507 PRELIMINARY; PRT; 334 AA.
AC O46507
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Plasminogen (Fragment).
GN BABPSPSG
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Buthera; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Cox L.A., Jett C., Hixson J.E.;
RT "Molecular Basis of the Apolipoprotein (a) Null Phenotype: A Splice
RT Site Mutation is Associated with Deletion of a Single Exon in a Null
RT Allele."
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL: AF029692; AAB97887.1; .
DR HSSP: P00747; SHGP.
DR MEROPS: S01.233; .
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR003966; Prothrombin.
DR Pfam: PF00051; Kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR PRINTS: PR00018; KRINGLE.
DR PRINTS: PR01505; PROTHROMBIN.
DR ProDom: PD000395; Kringle; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00021; KRINGLE 1; 1.
DR PROSITE: PS00070; KRINGLE 2; 1.
DR PROSITE: PS0240; TRYPSIN_DOM; 1.
DR PROSITE: PS0134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
SQ SEQUENCE 334 AA; 36791 MW; C7DC06B03B965286 CRC64;

Query Match 21.0%; Score 482.5; DB 6; Length 334;
Best Local Similarity 35.4%; Pred. No. 4.3e-38;
Matches 129; Conservative 46; Mismatches 134; Indels 55; Gaps 16;
QY 50 CYEGNGHYRGKASDTDMGRCLPWN SATVLQOYHAAH---RSDALQLGLGKHNYCRNP 105
Db 5 CMFGNGKRYGKATVTGTPQEWAA-----KEPHSLIFTPTTPRAGLEK-NYCRNP 58
QY 106 D-NRRRPWCYVQVGLKPLVQECWHDCA DGKPKSPPEELKFQCGQKTLRPR---FKIIG 161
Db 59 DGDVGGPWCYT-TNPRKLYDCVDPQASS-----SFDGKPKQVEPKKPCGRVVG 107

```

QY 162 GEFTIENQWPAALYRRHGGSVYVCGGSLISPCWVISAHCIFIDYPKKEDYIVVLGR 221
Db 108 GCVAHAHSWQVSL--RTRFG--MHECGGTLSPEWVLTAAHCLSKSPSPFYKILGA 163
QY 222 SRLNSNTQGMKFF--VENILHKDYSGADTLAHHNDTALLKIRSKGRCQAQPSRTIQTIC 279
Db 164 HQ-----EVRLEPHVQIEVSMFSEPAGA---DIALKLSSP-----AIIIDKVIAP 209
QY 280 LPSMYNDPQF-----GTCEITGFKENSTDYLYPEQLKMTVVKLISHRECOQPHYVSEV 335
Db 210 LPS-----PNYVADRTRCFITGWTGOT--YGAGLKEARLPVIEKVCNRYEFNGRV 263
QY 336 TTKMLCAADPOWTDSCQSGGSLVCSLQGRMTLTGIVSWGRCALDKDPGVYTRVSHF 395
Db 264 KSTELCAGHLAGGDSQSGGSLVCFEKDKYILQGVTSWGLGRCARPNKPGVYVRSRF 323
QY 396 LPWI 399
Db 324 VTWI 327

RESULT 27
Q18783
ID 018783 PRELIMINARY; PRT; 806 AA.
AC 018783;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen.
OS Macropus eugenii (Tamar wallaby).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.
OX NCBI_TaxID=9315;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98004511; PubMed=9342350;
RA Lawn R.M., Schwartz K., Patthy L.;
RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).
CC -/- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS.
DR EMBL; AF012297; AAB65760.1; -.
DR HSP; P00747; SHPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS00070; KRINGLE_2; 5.
DR PROSITE; PS0240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

```

Query Match 20.5%; Score 472.5; DB 6; Length 806;
 Best Local Similarity 32.0%; Pred. No. 1.2e-36;
 Matches 132; Conservative 43; Mismatches 162; Indels 75; Gaps 16;

```

QY 13 CLNGATCVSNKYFSNIHWCNCPKFKGGQHCIDSKTCYEGNGHPIYRQKASTDITMRPCL 72
Db 448 CSGTGSTVLNAQTRV-----PSVDTTSHFESD-----CMYSGKDYRGKRSTTITGLCQ 498
QY 73 PMSATVLQOTVHAH-----RSDALQLGLKHNVCNRPD--NRRRPWCYVQVGLKPLVQECMV 128
Db 499 AWTA-----QEPHRTITPTDYPRAGLEENVCNRPDGDGPNGPWCYT--TNPKKLEFDYCDI 552
QY 129 HDCADGKPKSSPEELKFKCGQKTLRPR---FKIIGGEFTTIENQWPAALYRRHGGSV 185
Db 553 PQCV-----FDCGKPRVEPOKCPGRIVGGCYAOPHSWPW--QISLRTFGE- 600
QY 186 TVCGGSLISPCWVISAHCIFIDYPKKEDYIVLGRSRLNSNTQGMKFEVENILHKDY 245
Db 601 -HFCGGTIAFQWLVLTAAHCLERSQWPGAYKVLG-----LHREV 639
QY 246 SADTLAHH-----NDIALLKIRSKGRCQAQPSRTITQICLPSMYNDPQGTSC 293
Db 640 NPESYSQIEGVSRLEFKGPLAADIALLKL-----NRPAINDKVIPACLFSDQPMVDRTLIC 695
QY 294 EITGFKENSTDYLYPE--OLKMTVVKLISHRECOQPHYVSEVTTKMLCAADPOWTDSC 352
Db 696 HVTGWGDTQGTG---PRGLLKQASLEVIDNRYCNRHEYLNGRVKSTELCAGHVGRCGDS 752
QY 353 QGDSGGLVCSLQGRMTLTGIVSWGRCALDKDPGVYTRVSHFPLWIRSHTK 404
Db 753 QGDSGGLVCFEDDKVYLGQVTSWGLGRCARPNKPGVYVRSYIISWIEDVMK 804

RESULT 28
Q9Y1V3
ID Q9Y1V3 PRELIMINARY; PRT; 868 AA.
AC Q9Y1V3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tunicate retinoic acid-inducible modular protease precursor.
GN TRAMP.
OS Polyandrocarpa misakiensis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Styelidae; Polyandrocarpa.
OX NCBI_TaxID=7723;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=white SPOT.
RX MEDLINE=99423646; PubMed=10491255;
RA Ohashi M., Kawamura K., Fujii N., Yubisui T., Fujiwara S.;
RT "A retinoic acid-inducible modular protease in budding ascidians.";
RL Dev. Biol. 214:38-45(1999).
CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AB030007; BAA82522.1; -.
DR HSP; P00763; IDPO.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF00057; ldl_recept_a; 3.
DR Pfam; PF00530; SRCR; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR PRINTS; PR00258; SPERACTRCPTR.
DR SMART; SM00192; LDLa; 3.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00202; SR_2.
DR SMART; SM00020; TRY_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; 3.
DR PROSITE; PS00068; LDLRA_2; 3.
DR PROSITE; PS0287; SRCR_2; 2.
DR PROSITE; PS0240; TRYPsin DOM; 1.
DR PROSITE; PS00134; TRYPsin HIS; 1.
DR PROSITE; PS00135; TRYPsin SER; 1.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 42.1951 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-3

Perfect score: 2301

Sequence: 1 SNEHQVPSNCDLNGTCTV.....VSHFLPWIRSHYKXENGLAL 411

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2301	100.0	411	10	US-09-880-503-3
2	2301	100.0	431	12	US-10-301-822-161
3	2301	100.0	431	12	US-10-131-985-21
4	2301	100.0	431	14	US-10-076-421-2
5	2301	100.0	431	15	US-10-171-311-184
6	2299	99.9	431	12	US-10-247-671-149
7	2298	99.9	431	15	US-10-193-656-4
8	2291	99.6	411	12	US-10-407-821-2
9	2283	99.2	431	10	US-09-264-468B-1
10	2243	97.5	403	10	US-09-880-503-6
11	1737	75.5	323	10	US-09-880-503-7
12	1535	66.7	337	15	US-10-106-698-6266
13	1508	65.5	276	10	US-09-880-503-5
14	1464	63.6	268	12	US-10-407-821-3
15	1333	57.9	246	10	US-09-264-468B-2

16	1318	57.3	241	11	US-09-898-837A-47	Sequence 47, Appl
17	864.5	37.6	527	11	US-09-987-457-18	Sequence 18, Appl
18	864.5	37.6	527	11	US-09-987-455-19	Sequence 19, Appl
19	864.5	37.6	562	9	US-09-969-271-7	Sequence 7, Appl
20	864.5	37.6	562	10	US-09-974-298-145	Sequence 145, App
21	864.5	37.6	562	12	US-10-443-701-4	Sequence 4, Appl
22	864.5	37.6	562	15	US-10-193-656-8	Sequence 8, Appl
23	837	36.4	143	10	US-09-880-503-8	Sequence 8, Appl
24	793	34.5	135	10	US-09-880-503-4	Sequence 4, Appl
25	793	34.5	138	15	US-09-984-186-12	Sequence 12, Appl
26	793	34.5	138	15	US-10-237-667-12	Sequence 12, Appl
27	793	34.5	138	15	US-10-237-708-12	Sequence 12, Appl
28	793	34.5	138	15	US-10-237-866-12	Sequence 12, Appl
29	793	34.5	138	15	US-10-237-871-12	Sequence 12, Appl
30	793	34.5	138	15	US-10-237-624-12	Sequence 12, Appl
31	785.5	34.1	372	9	US-09-084-491A-3	Sequence 3, Appl
32	785.5	34.1	372	14	US-10-102-704-3	Sequence 3, Appl
33	780	33.9	354	11	US-09-987-457-10	Sequence 10, Appl
34	780	33.9	354	11	US-09-987-455-11	Sequence 11, Appl
35	780	33.9	377	11	US-09-987-455-8	Sequence 8, Appl
36	741	32.2	343	11	US-09-987-457-14	Sequence 14, Appl
37	741	32.2	343	11	US-09-987-455-15	Sequence 15, Appl
38	740	32.2	339	11	US-09-987-457-12	Sequence 12, Appl
39	740	32.2	339	11	US-09-987-455-13	Sequence 13, Appl
40	737	32.0	331	11	US-09-987-457-11	Sequence 11, Appl
41	737	32.0	331	11	US-09-987-455-12	Sequence 12, Appl
42	737	32.0	343	11	US-09-987-457-15	Sequence 15, Appl
43	737	32.0	343	11	US-09-987-455-16	Sequence 16, Appl
44	735.5	32.0	655	15	US-10-172-712-28	Sequence 28, Appl
45	733	31.9	335	11	US-09-987-457-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match	100.0%	Score 2301;	DB 10;	Length 411;
Best Local Similarity	100.0%	Pred. No. 2.5e-194;		
Matches 411;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	SNEHQVPSNCDLNGTCTVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG	60	
DB	1	SNEHQVPSNCDLNGTCTVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG	60	
QY	61	KASTDTMGPCLPWNSATVLQOTYHAHRSALQGLGKKNYCRNPNRRPWCYVOVGLK	120	
DB	61	KASTDTMGPCLPWNSATVLQOTYHAHRSALQGLGKKNYCRNPNRRPWCYVOVGLK	120	
QY	121	PLVQECWHDCAADGKKPSPPEELKFCQCKTLRPRFKLIIGFEFTTIENQPFALYRRH	180	
DB	121	PLVQECWHDCAADGKKPSPPEELKFCQCKTLRPRFKLIIGFEFTTIENQPFALYRRH	180	

```
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
Db 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
QY 241 LHKDYSADTLAHNDIALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCIEITGFGK 300
Db 241 LHKDYSADTLAHNDIALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCIEITGFGK 300
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 360
Db 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 360
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411

RESULT 2
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-029PRNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 2301; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFWAAIYRRH 180
Db 141 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFWAAIYRRH 200
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCIEITGFGK 300
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCIEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 4
US-10-076-421-2
; Sequence 2, Application US/10076421
```

```
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 3
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Ocleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 100.0%; Score 2301; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 21 SNEHQVPSNCDCLNGTGVSNKYFSNIHWCNPKFGGQHCIDKSKTCYEGNGHYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNYCRNPONRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFWAAIYRRH 180
Db 141 PLVQECMVHDCADGKPKSPPEELKFCQCGOKTLRPRFKIIGGEFTTIENQPFWAAIYRRH 200
QY 181 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
Db 201 RGSVTVVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 260
QY 241 LHKDYSADTLAHNDIALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCIEITGFGK 300
Db 261 LHKDYSADTLAHNDIALLKIRSKGRCAPSRITQICLPSMYNDPQFTSCIEITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPEQLKMTVVKLIISHRECQOPHYGSEVTTKMLCAADPOWKTDSCQDSSGGL 380
QY 361 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 4
US-10-076-421-2
; Sequence 2, Application US/10076421
```

```
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAVAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match      100.0%; Score 2301; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 80
Qy 61 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
Qy 121 PLVOECMVHDCADGKPSPEELKFCQOGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 180
Db 141 PLVOECMVHDCADGKPSPEELKFCQOGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
Qy 181 RGSVTVVCGSLSPCWVISAHCFTIDYPKEDYIYVLSRLNSNTQGMKFEVENLI 240
Db 201 RGSVTVVCGSLSPCWVISAHCFTIDYPKEDYIYVLSRLNSNTQGMKFEVENLI 260
Qy 241 LHKDYSADTLAHNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFGTSCBITGFGK 300
Db 261 LHKDYSADTLAHNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFGTSCBITGFGK 320
Qy 301 ENSTDYLYPQLKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPQLKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 5
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159

; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAVAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match      100.0%; Score 2301; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.7e-194;
Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEDKSKTCYEGNGHFYRG 80
Qy 61 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASDTMGRPCLPWNSATVLTQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
Qy 121 PLVOECMVHDCADGKPSPEELKFCQOGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 180
Db 141 PLVOECMVHDCADGKPSPEELKFCQOGKTLRPRFKIIGGEFTTIENQPFPAIYRRH 200
Qy 181 RGSVTVVCGSLSPCWVISAHCFTIDYPKEDYIYVLSRLNSNTQGMKFEVENLI 240
Db 201 RGSVTVVCGSLSPCWVISAHCFTIDYPKEDYIYVLSRLNSNTQGMKFEVENLI 260
Qy 241 LHKDYSADTLAHNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFGTSCBITGFGK 300
Db 261 LHKDYSADTLAHNDIALKIRSKGRCAOPSRITQICLPSMYNDPQFGTSCBITGFGK 320
Qy 301 ENSTDYLYPQLKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 360
Db 321 ENSTDYLYPQLKMTVVKLISHRECQOPHYVGSVTTKMLCAADPQWKTDSCQDSSGGL 380
Qy 361 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 411
Db 381 VCSLQGRMTLTGIVSWGRCALDKPGVYTRVSHFLPWIRSHTKENGLAL 431

RESULT 6
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match      99.9%; Score 2299; DB 12; Length 431;
```


Best Local Similarity 99.8%; Pred. No. 4e-194; Matches 410; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCIEDKSKTCYEGNGHYVRG 60
Db	21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCIEDKSKTCYEGNGHYVRG 80
Qy	61 KASDTDTWGRPCLPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPDRNRPPWCVVQVGLK 120
Db	81 KASDTDTWGRPCLPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPDRNRPPWCVVQVGLK 140
Qy	121 PLVOECMVHDCADGKKSPPEELKFCQGQKTLRPRFKIIGGETTIENQWFAAIVRRH 180
Db	141 PLVOECMVHDCADGKKSPPEELKFCQGQKTLRPRFKIIGGETTIENQWFAAIVRRH 200
Qy	181 RGSVTVVCGGSLSPCWISATHCFTDYPKKEDYIVLGRSLNSNTQGEKMEVENLI 240
Db	201 RGSVTVVCGGSLSPCWISATHCFTDYPKKEDYIVLGRSLNSNTQGEKMEVENLI 260
Qy	241 LHKDYSADTLAHNDIALALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 300
Db	261 LHKDYSADTLAHNDIALALKIRSEGRCAQPSRTIQTICLPSMYNDPQFGTSCEITGFGK 320
Qy	301 ENSTDYLYPEQLKMTVVKLISHRECCQPHYYGSEVTTKMLCAADPQWKTDSCQDGSGLP 360
Db	321 ENSTDYLYPEQLKMTVVKLISHRECCQPHYYGSEVTTKMLCAADPQWKTDSCQDGSGLP 380
Qy	361 VCSLQGRMTLTGIVSWRGCCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
Db	381 VCSLQGRMTLTGIVSWRGCCALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 431

```

RESULT 7
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

```

```

Qy 121 PLVQECMVHDCADGKKSPSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 180
Db 141 PLVQECMVHDCADGKKSPSPPEELKFCQGCKTLRPRFKIIGGEFTTIENQWPFAAIYRRH 200
Qy 181 RGGSVTYVCGSGLISPCWVISATHCFIDYPKKEDYIVVLGSRSLNSNTQGMKFEVENLI 240
Db 201 RGGSVTYVCGSGLISPCWVISATHCFIDYPKKEDYIVVLGSRSLNSNTQGMKFEVENLI 260
Qy 241 LHKDYSADTLAHNDNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDPQGTSCETIGFGK 300
Db 261 LHKDYSADTLAHNDNDIALLKIRSKEGRCAQPSRTIQTICLPSMYNDPQGTSCETIGFGK 320
Qy 301 ENSTDYILYPLQKMTVVVKLISHRCCQPHYYGSEVTTKMLCAADPQWKTDSCQGSQGGPL 360
Db 321 ENSTDYILYPLQKMTVVVKLISHRCCQPHYYGSEVTTKMLCAADPQWKTDSCQGSQGGPL 380
Qy 361 VCSLQGRMTLTIGVSWGRGCAALKDKPGVYTRVSHFLPWIRSHTKKEENGAL 411
Db 381 VCSLQGRMTLTIGVSWGRGCAALKDKPGVYTRVSHFLPWIRSHTKKEENGAL 431

RESULT 8
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022UUS
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

```

Query Match	99.6%;	Score 2291;	DB 12;	Length 411;			
Best Local Similarity	99.8%;	Pred. No. 1.9e-193;					
Matches 410;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;			
QY	1	SNELHQPVSNCDC	CLNGGTCVSNKYF	SIHWCNPKF	GGOHCEIDKSKTCYEGNGH	FVRG 60	
Db	1	SNELHQPVSNCDC	CLNGGTCVSNKYF	SIHWCNPKF	GGOHCEIDKSKTCYEGNGH	FVRG 60	
QY	61	KASTDTMGRPCL	PWNSATVLQQT	YHAHRSDALQLGLG	KHNYCRNPNRNR	PPWCYQVGLK 120	
Db	61	KASTDTMGRPCL	PWNSATVLQQT	YHAHRSDALQLGLG	KHNYCRNPNRNR	PPWCYQVGLK 120	
QY	121	PLVQECVMHDC	ADGKXPSPEEL	KFCGQKTLRPR	KIIGCEFTTIENQ	PFAAIVRRH 180	
Db	121	LLVQECVMHDC	ADGKXPSPEEL	KFCGQKTLRPR	KIIGCEFTTIENQ	PFAAIVRRH 180	
QY	181	RGGSVTYVCGG	SLISPCWVLSATH	CFIDYPKEDYIVLGR	SLNSNTGEMKFE	VENLI 240	
Db	181	RGGSVTYVCGG	SLISPCWVLSATH	CFIDYPKEDYIVLGR	SLNSNTGEMKFE	VENLI 240	
QY	241	LHKDYSADTLA	HNHDIALLKIRSE	KGRCAQPSRTIQ	ITICLPSMYNDP	QFCTSGEITGFGK 300	
Db	241	LHKDYSADTLA	HNHDIALLKIRSE	KGRCAQPSRTIQ	ITICLPSMYNDP	QFCTSGEITGFGK 300	
QY	301	ENSTDLYLPEQL	KMTVVVKLISHRE	CQOPHYHGVSEV	TTKMLCAADPQ	MKTDCQDGS	GGPL 360
Db	301	ENSTDLYLPEQL	KMTVVVKLISHRE	CQOPHYHGVSEV	TTKMLCAADPQ	MKTDCQDGS	GGPL 360

QY 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411
Db 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411

RESULT 9

US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.PL
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US/09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 99.2%; Score 2283; DB 10; Length 431;
Best Local Similarity 99.5%; Pred. No. 1e-192;
Matches 409; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDLNGGTCVSNKYFNSIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDLNGGTCVSNKYFNSIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGKPKSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHH 180
Db 141 PLVQECMVHDCADGKPKSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHH 200
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGEKMFVENLI 240
Db 201 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGEKMFVENLI 260
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 300
Db 261 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 320
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGGFL 360
Db 321 EXSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGGFL 380
QY 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411

Db 381 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 431

RESULT 10

US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 97.5%; Score 2243; DB 10; Length 403;
Best Local Similarity 98.1%; Pred. No. 3.2e-189;
Matches 403; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 SNELHQPNSCDLNGGTCVSNKYFNSIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPNSCDLNGGTCVSNKYFNSIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRCPLPWSATVLOQTYHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPELKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHH 180
Db 121 PLVQECMVHDCADGK-----LKFQCGQKTLRPRFKIIGGEFTTIENQWFAAIYRHH 172
QY 181 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGEKMFVENLI 240
Db 173 RGSVTVYVCGSLISPCWVISATHCFIDYPPKEDYIVYLGSRSLNSNTQGEKMFVENLI 232
QY 241 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 300
Db 233 LHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPSMYNDPQGTSCETITGFGK 292
QY 301 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGGFL 360
Db 293 ENSTDYLYPEQLKMTVVKLIISHRECCQPHYVGSVTTKMLCAADPQWKTDSCQDGGGFL 352
QY 361 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 411
Db 353 VCSLQGRMTLTGIVSWGRGKALDKPGVYTRVSHFLPWIRSHTEENGLAL 403

RESULT 11

US-09-880-503-7
; Sequence 7, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20

; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-503-7

Query Match 75.5%; Score 1737; DB 10; Length 323;
 Best Local Similarity 78.6%; Pred. No. 8.8e-145;
 Matches 323; Conservative 0; Mismatches 0; Indels 88; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDSKTCYEGNGHFYRG 60
 DB 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDSKS-----47

QY 61 KASTDTMGRPCLPWNSATVLTQTYHAHRSDALQLGLGKKNYCRNPNRRPWCYVQVGLK 120
 DB -----47

QY 121 PLVQECMVHDCADGKKPSSPEELKFQCGQKTLRPRFKIIGGFTTIENQWFAAIYRRH 180
 DB -----KSSPPEELKFQCGQKTLRPRFKIIGGFTTIENQWFAAIYRRH 92

QY 181 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 240
 DB 93 RGSVTVVCGGSLISPCWISATHCFIDYPKKEDYIVYLGRSLNSNTQGMKFEVENLI 152

QY 241 LHKDYSADTLAHHNDIALLKIRSKGRCAPQSRITQICLPSMYNDPQFGTSCETGFGK 300
 DB 153 LHKDYSADTLAHHNDIALLKIRSKGRCAPQSRITQICLPSMYNDPQFGTSCETGFGK 212

QY 301 ENSTDLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPL 360
 DB 213 ENSTDLYPEQLKMTVVKLISHRECCQPHYVYGVSEVTTKMLCAADPQWKTDSCQDGGPL 272

QY 361 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 411
 DB 273 VCSLQGRMTLTGIVSWGRGALKDKPGVYTRVSHFLPWIRSHTKKEENGLAL 323

RESULT 12
 US-10-106-698-6266
 ; Sequence 6266, Application US/10106698
 ; Publication No. US20030109690A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
 ; FILE REFERENCE: PA005P1
 ; CURRENT APPLICATION NUMBER: US/10/106,698
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: PCT/US00/26524
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: US 60/157,137
 ; PRIOR FILING DATE: 1999-09-29
 ; PRIOR APPLICATION NUMBER: US 60/163,280
 ; PRIOR FILING DATE: 1999-11-03
 ; NUMBER OF SEQ ID NOS: 8564
 ; SOFTWARE: PatentIn Ver. 3.0
 ; SEQ ID NO 6266
 ; LENGTH: 337
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-106-698-6266

Query Match 66.7%; Score 1535; DB 15; Length 337;
 Best Local Similarity 95.2%; Pred. No. 5.7e-127;
 Matches 275; Conservative 3; Mismatches 5; Indels 6; Gaps 1;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDSKTCYEGNGHFYRG 60
 DB 27 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFGGQHCIDSKTCYEGNGHFYRG 86

```
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-3

Query Match
Best Local Similarity 100.0%; Pred. No. 7.6e-121; Length 268;
Matches 268; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

144 LKFCGQKTLRPRFKIIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISAT 203
Db 1 LKFCGQKTLRPRFKIIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISAT 60

204 HCFIDYPKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHNDIALKIRS 263
Db 61 HCFIDYPKEDYIVYLGSRSLNSNTQGMKPEVENLILHKDYSADTLAHNDIALKIRS 120

264 KEGRCAQPSRTIQTICLPMSYNDPOFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHR 323
Db 121 KEGRCAQPSRTIQTICLPMSYNDPOFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHR 180

324 ECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALK 383
Db 181 ECOQPHYGVSEVTTKMLCAADPOWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALK 240

384 DKPGVTVRVSHFLPWIRSHTKENGLAL 411
Db 241 DKPGVTVRVSHFLPWIRSHTKENGLAL 268

RESULT 15
US-09-264-468B-2
; Sequence 2, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Sevetin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-264-468B-2

Query Match
Best Local Similarity 57.9%; Score 1333; DB 10; Length 246;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

159 IIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVY 218
QY 159 IIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVY 218
```

```
Db 1 IIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVY 60
QY 219 LGRSLNSNTQGMKPEVENLILHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQT 278
Db 61 LGRSLNSNTQGMKPEVENLILHKDYSADTLAHNDIALKIRSKGRCACQPSRTIQT 120
QY 279 CLPSMYNDPOFGTSCEITGFGKENSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTK 338
Db 121 ALPSMYNDPOFGTSCEITGFGKEQSTDYLYPEQLKMTVVKLISHRECOQPHYGVSEVTTK 180
QY 339 MLCADPOWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPW 398
Db 181 MLCADPOWKTDSCQDGGPLVCSLQGRMTLTGIVSWGRCALKDKPGVTVRVSHFLPW 240
QY 399 IRSHTK 404
Db 241 IRSHTK 246

RESULT 16
US-09-898-837A-47
; Sequence 47, Application US/09898837A
; Publication No. US20030077697A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Herrmann, John L.
; APPLICANT: Burgess, Catherine
; APPLICANT: Fernandes, Elma
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Rastelli, Luca
; APPLICANT: CuraGen Corporation
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: MacDougall, John R.
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
; FILE REFERENCE: 15966-598 CIP
; CURRENT APPLICATION NUMBER: US/09/898,837A
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 241
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-898-837A-47

Query Match
Best Local Similarity 57.3%; Score 1318; DB 11; Length 241;
Matches 240; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

159 IIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVY 218
QY 159 IIGGFTTIENQFWFAAIYRRHGGSVTVVCGSLISPCWVISATHCFIDYPKEDYIVY 218
```

Db 1 IIGGEFTIENQPMFAAIYRRHGGSVTYVCGGSLMSPCWISATHCFIDYPPKEDIVY 60
QY 219 LGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQT 278
Db 61 LGRSRLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQT 120
QY 279 CLPSMYNDPQFGTSCEITGFGKENSTLYLPEQLKMTVVKLISHRECOQPHYGVSEVTTK 338
Db 121 CLPSMYNDPQFGTSCEITGFGKENSTLYLPEQLKMTVVKLISHRECOQPHYGVSEVTTK 180
QY 339 MLCAADPQWKTDCSQGDSGGPLVCSLQGRMTLTGIVSGWRCALKDKPGVTVRVSHFLPW 398
Db 181 MLCAADPQWKTDCSQGDSGGPLVCSLQGRMTLTGIVSGWRCALKDKPGVTVRVSHFLPW 240
QY 399 I 399
Db 241 I 241

RESULT 17
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.2180001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens (tPA)
US-09-987-457-18

Query Match 37.6%; Score 864.5; DB 11; Length 527;
Best Local Similarity 38.1%; Pred. No. 1.1e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWNCNPKKFGQHCIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPCFNGGTCQQALYFSDP-VCQCPGFGAGKCCIEDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVG 118
Db 101 RGTWSTAESAECTNMNSALAQPYSGRRPDAILRLGLGNHNYCRNPNDRDSKPCWYVFKA 160
QY 119 LKPLVQCMVHDCADG-----KGPS 138
Db 161 GKYSEFCSTPACSEGNSDCYFNGSAYRGTHSLTESGASCLPWNSMILIGKYVTAQNP 220
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db 221 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQPOQPRIKGGL 280
QY 164 FTTIENQPMFAAIYRRH-RGGSVTVYVCGGSLISPCWISATHCFIDYPPKEDIVYVGLRS 222
Db 281 FADIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAACFQERPPPHLTVILGRT 340
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPS 282
Db 341 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDNIALQLKSDSSRCAQESSVVRTVCLPP 398
QY 283 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPS 282
Db 341 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDNIALQLKSDSSRCAQESSVVRTVCLPP 398

QY 283 MYNDPQFGTSCEITGFGKENSTLYLPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 342
Db 399 ADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPPSSRCTSQHLLNRTVTDNMLCA 458
QY 343 AD-----PQWKT-DSCQDGGGGLVCSLQGRMTLTGIVSGWRCALKDKPGVTVRVSHFL 396
Db 459 GDTSGGFPQANLHDACQDGGGGLVCLNDGRMTLVGIISWGLGCGQKDVPGVYTKVTNYL 518
QY 397 PWIRSHTK 404
Db 519 DWIRDNMR 526
RESULT 18
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.2190001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 37.6%; Score 864.5; DB 11; Length 527;
Best Local Similarity 38.1%; Pred. No. 1.1e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;
QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWNCNPKKFGQHCIDKSKTCYEGNGHFY 58
Db 42 QCHSVPVKSCSEPCFNGGTCQQALYFSDP-VCQCPGFGAGKCCIEDTRATCYEDQGISY 100
QY 59 RGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVG 118
Db 101 RGTWSTAESAECTNMNSALAQPYSGRRPDAILRLGLGNHNYCRNPNDRDSKPCWYVFKA 160
QY 119 LKPLVQCMVHDCADG-----KGPS 138
Db 161 GKYSEFCSTPACSEGNSDCYFNGSAYRGTHSLTESGASCLPWNSMILIGKYVTAQNP 220
QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db 221 AQALGLGKHNYCRNPDGAKPWCHLVKNRRLTWECYDVPSCSTGLRQYQPOQPRIKGGL 280
QY 164 FTTIENQPMFAAIYRRH-RGGSVTVYVCGGSLISPCWISATHCFIDYPPKEDIVYVGLRS 222
Db 281 FADIASHPWQAIFAKHRRSPGERFLCGGILISSCWILSAACFQERPPPHLTVILGRT 340
QY 223 RLNSNTQGMKFEVENLILHKDYSADTLAHNDIALKIRSKGRCQAQPSRTIQTICLPS 282
Db 341 YRVVPGEEQKFEVEKYIVHKEFDDDT--YNDNIALQLKSDSSRCAQESSVVRTVCLPP 398
QY 283 MYNDPQFGTSCEITGFGKENSTLYLPEQLKMTVVKLISHRECOQPHYGVSEVTTKMLCA 342
Db 399 ADLQLPDWTCELSGKGHEALSPFYSERLKEAHVRLYPPSSRCTSQHLLNRTVTDNMLCA 458

```
QY 343 AD-----POWKT-DSQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 459 GDTSGGPGQANLHDA CQDGGGGLVCLNDGRMTLVGLIISWGLCGGQKDPGVVTKVTNYL 518

QY 397 PWIRSHK 404
Db 519 DWIRDNMR 526

RESULT 19
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 37.6%; Score 864.5; DB 9; Length 562;
Best Local Similarity 38.1%; Pred. No. 1.2e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNSIHWNCPCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCNGGTCQOALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLQOTVHAHRS DALQLGLGKHNYCRNPDRRRRWCYVQVG 118
Db 136 RGTWSTAESGAECTNNSSALAKPISGRPRDAIRLGLGNHNYCRNPDRDRSKPWCYVFKA 195

QY 119 LKPLVQECMVHDCADG-----KPKS 138
Db 196 GKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVYTAQNP 255

QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db 256 AQALGLGKKNYCRNPDPDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPQFRKXGGL 315

QY 164 FTTIENQPFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 222
Db 316 FADIASHPPQAAIFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPFHHLTVILGRT 375

QY 223 RLNSNTQGMKEVENILHKDYSATLAHNDIALLKIRSKGRCAQPSRTIQTICLPS 282
Db 376 YRVVPGEERQKFEVEKYIVHKEPDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433

QY 283 MYNDPQFGTSCETGFGKENSDYLYPEQLKMTVVKLISHRECQOQPHYGSEVTTMLCA 342
Db 434 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCA 493

QY 343 AD-----POWKT-DSQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 494 GDTSGGPGQANLHDA CQDGGGGLVCLNDGRMTLVGLIISWGLCGGQKDPGVVTKVTNYL 553

QY 397 PWIRSHK 404
Db 554 DWIRDNMR 561
```

RESULT 20
US-09-974-298-145

```
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 145
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 1001470CD1
US-09-974-298-145

Query Match 37.6%; Score 864.5; DB 10; Length 562;
Best Local Similarity 38.1%; Pred. No. 1.2e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFNSIHWNCPCPKFGGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCNGGTCQOALYFSDP-VQCQPEGFAGKCCIEDTRATCYEDQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLQOTVHAHRS DALQLGLGKHNYCRNPDRRRRWCYVQVG 118
Db 136 RGTWSTAESGAECTNNSSALAKPISGRPRDAIRLGLGNHNYCRNPDRDRSKPWCYVFKA 195

QY 119 LKPLVQECMVHDCADG-----KPKS 138
Db 196 GKYSSEFCSTPACSEGNDCYFGNGSAYRGTHSLTESGASCLRWNSMILIGKVYTAQNP 255

QY 139 -----SPPEELK-----FOCG-QKTLRPRFKIIGGE 163
Db 256 AQALGLGKKNYCRNPDPDAKPWCHVLKNRRLTWECYCDVPSCSTCGLRQYSQPQFRKXGGL 315

QY 164 FTTIENQPFAAIYRRH-RGGSVTVYVCGGSLISPCWVISATHCFIDYPKKEDYIVYLGRS 222
Db 316 FADIASHPPQAAIFAKHRRSPGERFLCGGILISSCWILSAHCFQERFPFHHLTVILGRT 375

QY 223 RLNSNTQGMKEVENILHKDYSATLAHNDIALLKIRSKGRCAQPSRTIQTICLPS 282
Db 376 YRVVPGEERQKFEVEKYIVHKEPDDDT--YDNDIALQLKSDSSRCAQESSVVRTVCLPP 433

QY 283 MYNDPQFGTSCETGFGKENSDYLYPEQLKMTVVKLISHRECQOQPHYGSEVTTMLCA 342
Db 434 ADLQLPDWTECELSGYGKHEALSPFYSERLKEAHLVLPSSRCTSQHLLNRTVTDNMLCA 493

QY 343 AD-----POWKT-DSQGDGSGPLVCSLQGRMTLTGIVSWGRCALKDKPGVYTRVSHFL 396
Db 494 GDTSGGPGQANLHDA CQDGGGGLVCLNDGRMTLVGLIISWGLCGGQKDPGVVTKVTNYL 553

QY 397 PWIRSHK 404
Db 554 DWIRDNMR 561

RESULT 21
US-10-443-701-4
; Sequence 4, Application US/10443701
; Publication No. US20030199016A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Yuan
; TITLE OF INVENTION: REVERSE-PHASE HPLC ASSAY FOR PLASMINOGEN ACTIVATORS
; FILE REFERENCE: P1788R1
; CURRENT APPLICATION NUMBER: US/10/443,701
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: US/09/703,695
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00750
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: {1}..(562)
US-10-193-656-8

Query Match          37.6%; Score 864.5; DB 15; Length 562;
Best Local Similarity 38.4%; Pred. No. 1.2e-67;
Matches 186; Conservative 56; Mismatches 157; Indels 89; Gaps 11;

Qy      3 ELHQVP--SNCD-----CLNGGTCVSNKPYFSNIHWCNCPKFGGQHCEIDKSKTCEYEGNHFY 58
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      77 QCHSVPVKSCSEPCFNGGTCCQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 135

Qy      59 RGKASDTDMGRPCLPWNSATVLQOTYAHRSDALQLGLGHKNYCRPNDRRRRWCVYQVG 118
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      136 RGTWSTAESAECTNNWNSSALAOKPYSGRRPDAILRLGLGNHNHYCRNPDRDKPCWCVYKA 195

Qy      119 LKPLVOBCMVDHCADG-----FCNGSAYRGTHSLTESGASCLPWSMILIGKVYTAQNPS 138
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      196 GKYSFECSPTACSEGNSDCYFCNGSAYRGTHSLTESGASCLPWSMILIGKVYTAQNPS 255

Qy      139 -----SPPELK-----FOCG-QKTLRPFKKIIGE 163
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      256 AQALGLGKHNYCRNPDGDAPCWCHLVKNRRLTWECYCDVPSCSTCGLRQYSQPQFRINGKGL 315

Qy      164 FTTIENQPFAAIYRRH-RGGSVTYYCGGSLISPCMWISATHCFIDYPKKEDIYVILGRS 222
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      316 FADIASHPWQAALFAHRHSFGERFLCGGLISSCMILSAACHFQERFPHPHLTVLIGRT 375

Qy      223 RLNSNTQGENMKFEVENILHKDYASDTLAHNNDIALIKIRSKBRCACAQPSRTIQTICLPSS 282
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      376 YRVVPGEEEQKFEVEKYIVHKFDDDT--YDNDIALIQLKSDSSRCAQESSVVRTVCPLPP 433

Qy      283 MYNDPQGTSCETITFGKENSTDYLIPQELKMVTVKLI SHRECOOPHYGVSEVTTMLCA 342
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      434 ADLQLPDWTICELSGYGKEALSPFYSERUKEAHRVLYPSSRCTQSOLLNARTVTDNMICA 493

Qy      343 AD-----POWK--DSGCGDSGGPLVCSLOGRMTLTGITVSMGRGCALKDKPGVYTRVSHFL 396
        : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db      494 GDTRSFGGPQANLHDACCDSGGPLVCLNDGMRTVLVGIISWLGCGCKXDVPGVYTKVTNYL 553

Qy      397 PWIRSHTK 404
        : ||| :
Db      554 DWIRDNR 561

```

US-09-880-503-8
 ; Sequence 8, Application US/09880503
 ; Patent No. US20020131964A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CINES, Douglas B
 ; APPLICANT: HIGAZI, Abd Al-Roof
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
 ; TITLE OF INVENTION: TISSUE CONTRACTABILITY
 ; FILE REFERENCE: 9596-331
 ; CURRENT APPLICATION NUMBER: US/09/880,503
 ; CURRENT FILING DATE: 2001-06-13
 ; PRIOR APPLICATION NUMBER: US 60/212,847
 ; PRIOR FILING DATE: 2000-06-20
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 143
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-880-503-8

Query Match 36.4%; Score 837; DB 10; Length 143;
 Best Local Similarity 100.0%; Pred No. 5, 8e-66.

Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDSKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDSKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWSNATVLQOQTYHAHRSDALQGLGKHNKNCNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWSNATVLQOQTYHAHRSDALQGLGKHNKNCNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGKPKSPPEE 143
DB 121 PLVQECMVHDCADGKPKSPPEE 143

RESULT 24
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match 34.5%; Score 793; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 4e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDSKSKTCYEGNGHFYRG 60
DB 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDSKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWSNATVLQOQTYHAHRSDALQGLGKHNKNCNPNRRPWCYVQVGLK 120
DB 61 KASTDTMGPRCLPWSNATVLQOQTYHAHRSDALQGLGKHNKNCNPNRRPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 25
US-09-884-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-884-186-12

Query Match 34.5%; Score 793; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDSKSKTCYEGNGHFYRG 60
DB 4 SNELHQPNSCDCLNGGTCVSNKYFSNIHWNCNPKKFGGQHCHIDSKSKTCYEGNGHFYRG 63
QY 61 KASTDTMGPRCLPWSNATVLQOQTYHAHRSDALQGLGKHNKNCNPNRRPWCYVQVGLK 120
DB 64 KASTDTMGPRCLPWSNATVLQOQTYHAHRSDALQGLGKHNKNCNPNRRPWCYVQVGLK 123
QY 121 PLVQECMVHDCADGK 135
DB 124 PLVQECMVHDCADGK 138

RESULT 26
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA


```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: System 7.1
/ SOFTWARE: Word 5.1 (Patentin)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/237,866
/ FILING DATE: 10-Sep-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/797,689
/ FILING DATE: 31-JAN-1997
/ APPLICATION NUMBER: US 08/256,927
/ FILING DATE: 28-JUL-1994
/ APPLICATION NUMBER: FR 92/01064
/ FILING DATE: 31-JAN-1992
/ APPLICATION NUMBER: PCI/FR93/00085
/ FILING DATE: 28-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith Ph.D., Julie K.
/ REGISTRATION NUMBER: P-38,619
/ REFERENCE/DOCKET NUMBER: ST92006-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (610) 454-3839
/ TELEFAX: (610) 454-3808
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 138 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12

Query Match 34.5%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 60
Db 4 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 63
QY 61 KASTDTMGPRCLPWNSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWNSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 123
QY 121 PLVQECWVHDCADGK 135
Db 124 PLVQECWVHDCADGK 138

RESULT 29
US-10-237-871-12
/ Sequence 12, Application US/10237871
/ Publication No. US20030036172A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleer, Reinhard
/ Fournier, Alain
/ Guitton, Jean-Dominique
/ Jung, Gerard
/ Yeh, Patrice
/ TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
/ PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
/ CONTAINING SAID POLYPEPTIDES
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Road, 3C43
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19426
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/

MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: System 7.1
/ SOFTWARE: Word 5.1 (Patentin)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/237,871
/ FILING DATE: 10-Sep-2002
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/797,689
/ FILING DATE: 31-JAN-1997
/ APPLICATION NUMBER: US 08/256,927
/ FILING DATE: 28-JUL-1994
/ APPLICATION NUMBER: FR 92/01064
/ FILING DATE: 31-JAN-1992
/ APPLICATION NUMBER: PCI/FR93/00085
/ FILING DATE: 28-JAN-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Smith Ph.D., Julie K.
/ REGISTRATION NUMBER: P-38,619
/ REFERENCE/DOCKET NUMBER: ST92006-US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (610) 454-3839
/ TELEFAX: (610) 454-3808
/ INFORMATION FOR SEQ ID NO: 12:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 138 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12

Query Match 34.5%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 60
Db 4 SNELHVPNSCDCLNGTGVSNKYFNSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFGYRG 63
QY 61 KASTDTMGPRCLPWNSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 120
Db 64 KASTDTMGPRCLPWNSATVLQQTYHAHRSDALQGLGKHNKNCNPNRRRPWCYVQVGLK 123
QY 121 PLVQECWVHDCADGK 135
Db 124 PLVQECWVHDCADGK 138

RESULT 30
US-10-237-624-12
/ Sequence 12, Application US/10237624
/ Publication No. US20030082747A1
/ GENERAL INFORMATION:
/ APPLICANT: Fleer, Reinhard
/ Fournier, Alain
/ Guitton, Jean-Dominique
/ Jung, Gerard
/ Yeh, Patrice
/ TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
/ PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
/ CONTAINING SAID POLYPEPTIDES
/ NUMBER OF SEQUENCES: 36
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Rhone-Poulenc Rorer Inc.
/ STREET: 500 Arcola Road, 3C43
/ CITY: Collegeville
/ STATE: PA
/ COUNTRY: USA
/ ZIP: 19426
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: Macintosh
/
```

OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (patentIn)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/237,624
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12

Query Match 34.5%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.1e-62;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
Qy 61 KASDTWTGRPCLPWNSATVLOOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 120
Db 64 KASDTWTGRPCLPWNSATVLOOTYHAHRSDALQLGLGKHNYCRNPNRRRPWCYVQVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

Search completed: December 3, 2003, 15:05:43
Job time : 44.1951 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 6.98276 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-4

Perfect score: 793

Sequence: 1 SNEHLQVPSNCDCLNGTCV.....QVGLKPLVQECMVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	138	2	US-08-797-689-12
2	793	100.0	200	4	US-09-101-272G-73
3	793	100.0	208	4	US-09-101-272G-98
4	793	100.0	411	1	US-08-087-163-1
5	793	100.0	411	1	US-08-286-748B-18
6	793	100.0	411	1	US-08-153-799-18
7	793	100.0	430	1	US-07-942-157A-3
8	793	100.0	430	6	5219569-2
9	793	100.0	431	4	US-08-101-272G-1
10	793	100.0	431	6	518829-1
11	788	99.4	194	4	US-09-101-272G-80
12	788	99.4	201	4	US-09-101-272G-96
13	783	98.7	411	3	US-09-181-816-1
14	782	98.6	411	2	US-08-560-098A-48
15	769	97.0	157	3	US-08-142-590B-25
16	515	64.9	432	2	US-08-560-098A-47
17	514	64.8	365	1	US-08-093-741-83
18	514	64.8	365	1	US-08-720-012-83
19	514	64.8	393	2	US-08-560-098A-44
20	514	64.8	393	3	US-08-967-024C-24
21	514	64.8	393	3	US-08-967-024C-25
22	513	64.7	89	4	US-09-101-272G-62
23	335.5	42.3	477	2	US-08-560-098A-51
24	328.5	41.4	527	1	US-07-609-510B-16
25	328.5	41.4	527	2	US-08-811-949-39
26	328.5	41.4	527	5	PCT-US91-01025A-2
27	328.5	41.4	527	6	5185259-8

28	328.5	41.4	527	6	5520913-1
29	328.5	41.4	546	6	5200340-6
30	328.5	41.4	562	2	US-08-811-949-43
31	328.5	41.4	562	2	US-08-560-098A-50
32	328.5	41.4	562	2	US-08-883-795A-38
33	328.5	41.4	562	6	5185259-3
34	328.5	41.4	562	6	5200340-2
35	328.5	41.4	562	6	5344773-2
36	321.5	40.5	562	6	5244676-5
37	300	37.8	233	3	US-08-438-745-15
38	300	37.8	233	3	US-08-438-745-17
39	300	37.8	233	3	US-09-219-019-15
40	300	37.8	233	3	US-09-219-019-17
41	300	37.8	233	5	PCT-US94-05669A-15
42	300	37.8	233	5	PCT-US94-05669A-17
43	300	37.8	235	3	US-08-438-745-13
44	300	37.8	235	3	US-09-219-019-13
45	300	37.8	235	5	PCT-US94-05669A-13

Patent No. 5520913
Patent No. 5200340
Sequence 43, Appl
Sequence 50, Appl
Sequence 38, Appl
Patent No. 5185259
Patent No. 5200340
Patent No. 5344773
Patent No. 5244676
Sequence 15, Appl
Sequence 17, Appl
Sequence 15, Appl
Sequence 17, Appl
Sequence 15, Appl
Sequence 17, Appl
Sequence 13, Appl
Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5876369
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guittion, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```
; MOLECULE TYPE: protein
US-08-797-689-12

Query Match      100.0%; Score 793; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63

Qy 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 64 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123

Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138

RESULT 2
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match      100.0%; Score 793; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.4e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 80

Qy 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155

RESULT 3
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101.272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATFHI-ML chimeric protein
US-09-101-272G-98

Query Match      100.0%; Score 793; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.6e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 2 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 61

Qy 61 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 62 KASDTMTGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 121

Qy 121 PLVQECMVHDCADGK 135
Db 122 PLVQECMVHDCADGK 136

RESULT 4
US-08-087-163-1
; Sequence 1, Application US/08087163
; Patent No. 5472692
; GENERAL INFORMATION:
; APPLICANT: Liu, Jian-Ning
; APPLICANT: Gurewich, Victor
; TITLE OF INVENTION: PRO-UKONINASE MUTANTS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,163
; FILING DATE: 07/02/93
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, J. Peter
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04353/003001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-087-163-1

Query Match      100.0%; Score 793; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB J SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASTDTMGRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
RESULT 5
US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-286-748B-18
Query Match 100.0%; Score 793; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASTDTMGRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

DB 121 PLVQECMVHDCADGK 135
RESULT 6
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07974
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-153-799-18
Query Match 100.0%; Score 793; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 9.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
DB 61 KASTDTMGRCLPWSATVLQOQTYHAHRSDALQGLGKHNYCRNPNRRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135


```

RESULT 7
US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A
; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "MAP signal"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
US-07-942-157A-3

```

```

Query Match 100.0%; Score 793; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 1e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 20 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 120
Db 80 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 139
QY 121 PLVQECMVHDCADGK 135
Db 140 PLVQECMVHDCADGK 154

```

```

RESULT 8
5219569-2
; Patent No. 5219569

```

```

; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR, GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/766,858
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
; 5219569-2

Query Match 100.0%; Score 793; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 1e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 140
QY 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155

```

```

RESULT 9
US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
US-09-101-272G-1

```

```

Query Match 100.0%; Score 793; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPWCYVOVGLK 140
QY 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155

```

```
RESULT 10
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO:1:
; LENGTH: 431
5188829-1
Query Match 100.0%; Score 793; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 1e-71;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 21 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 80
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
Db 81 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155
RESULT 11
US-09-101-272G-80
; Sequence 80, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; FILE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: O50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 80:
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI chimeric protein
US-09-101-272G-80
Query Match 99.4%; Score 788; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 1.3e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 2 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 61
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 121
QY 121 PLVQECMVHDCADG 134
Db 122 PLVQECMVHDCADG 135
RESULT 12
US-09-101-272G-96
; Sequence 96, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: O50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; PRIOR FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 96:
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATPHI-CL chimeric protein
US-09-101-272G-96
Query Match 99.4%; Score 788; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 1.4e-71;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 2 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 61
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
Db 62 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 121
QY 121 PLVQECMVHDCADG 134
Db 122 PLVQECMVHDCADG 135
RESULT 13
US-09-181-816-1
; Sequence 1, Application US/09181816
; Patent No. 6277818
; GENERAL INFORMATION:
; APPLICANT: MAZAR, Andrew P.
; APPLICANT: JONES, Terence R.
; TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
; TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
; FILE REFERENCE: 329042000300 SIDN 1-7
; CURRENT APPLICATION NUMBER: US/09/181,816
; CURRENT FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1:
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-181-816-1
Query Match 98.7%; Score 783; DB 3; Length 411;
Best Local Similarity 99.3%; Pred. No. 1e-70;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
Db 1 SNELHQPSPNCDCCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135
```

```
RESULT 14
US-08-560-098A-48
; Sequence 48, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-48
Query Match 98.6%; Score 782; DB 2; Length 411;
Best Local Similarity 99.3%; Pred. No. 1.3e-70;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFG 60
Db 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFG 60

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 15
US-08-142-590B-25
; Sequence 25, Application US/08142590B
; Patent No. 6120765
; GENERAL INFORMATION:
; APPLICANT: HIBINO, Tashihiko, Takahashi, Tadahito, HORII, Izumi, and
; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/142,590B
; FILING DATE: 25-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,318
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul L.
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: MGP-009CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 157 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-142-590B-25
Query Match 97.0%; Score 769; DB 3; Length 157;
Best Local Similarity 97.0%; Pred. No. 8.5e-70;
Matches 131; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFG 60
Db 1 SNELHQPNSCDCLNGTCTVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFG 60

QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120
Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNCRPNRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 16
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
```

```
/ APPLICATION NUMBER: US/08/560,098A
/ FILING DATE: 17-NOV-1995
/ PRIOR APPLICATION DATA: P 44 40 892.7
/ APPLICATION NUMBER: P 44 40 892.7
/ FILING DATE: 17-NOV-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EVANS, Joseph D.
/ REGISTRATION NUMBER: 26,269
/ REFERENCE/DOCKET NUMBER: 148/42448
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-8800
/ TELEFAX: (202) 628-8844
/ INFORMATION FOR SEQ ID NO: 47:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 432 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-560-098A-47

Query Match 64.9%; Score 515; DB 2; Length 432;
Best Local Similarity 75.8%; Pred. No. 7.6e-44;
Matches 93; Conservative 8; Mismatches 16; Indels 6; Gaps 1;

QY 13 CLNGGTCVSNKYPFNIHWCNCPKFGQHCCEIDKSKTCYEGNGHFGYRGKASTDTMGRPCL 72
Db 40 CVTGEFTPESHNGDPEIPEY-----LQISKTCYEGNGHFGYRGKASTDTMGRPCL 93
QY 73 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 132
Db 94 PWSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLKPLVQECMVHDC 153
QY 133 DGK 135
Db 154 DGK 156

RESULT 17
US-08-093-741-83
/ Sequence 83, Application US/08093741
/ Patent No. 5681721
/ GENERAL INFORMATION:
/ APPLICANT: STEFFENS, Gerd J.
/ APPLICANT: WENNDT, STEPHAN
/ APPLICANT: SCHNEIDER, JOHANNES
/ APPLICANT: HEINZEL-WIELAND, REGINA
/ APPLICANT: SAUNDERS, DEREK J.
/ TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
/ TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
/ NUMBER OF SEQUENCES: 83
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
/ STREET: 1200 G Street, N. W. Suite 700
/ CITY: Washington, D.C.
/ COUNTRY: U.S.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/093,741
/ FILING DATE: 20-JUL-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/093,741
/ FILING DATE: 20-JUL-1993
/ APPLICATION NUMBER: DE P43 23 754.1
/ FILING DATE: 15-JUL-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EVANS, Joseph D.
/ REGISTRATION NUMBER: 26,269
/ REFERENCE/DOCKET NUMBER: 148/41345
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202)628-8800
/ TELEFAX: (202)628-8844
/ INFORMATION FOR SEQ ID NO: 83:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 365 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ US-08-720-012-83

Query Match 64.8%; Score 514; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 7.9e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFGYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFGYRGKASTDTMGRPCLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPD 60
QY 107 NRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 NRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 18
US-08-720-012-83
/ Sequence 83, Application US/08720012
/ Patent No. 5747291
/ GENERAL INFORMATION:
/ APPLICANT: STEFFENS, Gerd J.
/ APPLICANT: WENNDT, STEPHAN
/ APPLICANT: SCHNEIDER, JOHANNES
/ APPLICANT: HEINZEL-WIELAND, REGINA
/ APPLICANT: SAUNDERS, DEREK J.
/ TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
/ TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
/ NUMBER OF SEQUENCES: 83
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
/ STREET: 1200 G Street, N. W. Suite 700
/ CITY: Washington, D.C.
/ COUNTRY: U.S.
/ ZIP: 20005
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/720,012
/ FILING DATE: 27-SEP-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/093,741
/ FILING DATE: 20-JUL-1993
/ APPLICATION NUMBER: DE P43 23 754.1
/ FILING DATE: 15-JUL-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: EVANS, Joseph D.
/ REGISTRATION NUMBER: 26,269
/ REFERENCE/DOCKET NUMBER: 148/41345
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202)628-8800
/ TELEFAX: (202)628-8844
/ INFORMATION FOR SEQ ID NO: 83:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 365 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ US-08-720-012-83

Query Match 64.8%; Score 514; DB 1; Length 365;
```

```
Best Local Similarity 100.0%; Pred. No. 7.9e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 1 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 60
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 NRRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 19
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-Inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-560-098A-44

Query Match 64.8%; Score 514; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.6e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 20
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 42 665.8
; FILING DATE: 30-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42444
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-967-024C-24

Query Match 64.8%; Score 514; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.6e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 106
Db 2 SKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPD 61
QY 107 NRRRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 62 NRRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 21
US-08-967-024C-25
; Sequence 25, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WNEBDT, Stephan
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSHA, Elke
; APPLICANT: HEINZEL-WIELAND, Regina
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 42 665.8
FILING DATE: 30-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42444
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 64.8%; Score 514; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 8.6e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 SKTCYEGNGHYFGKASTDTMGPRCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPD 106
DB 2 SKTCYEGNGHYFGKASTDTMGPRCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPD 61
QY 107 NRRPWCYVQVGLKPLVQECWVHDCADGX 135
DB 62 NRRPWCYVQVGLKPLVQECWVHDCADGX 90

RESULT 22
US-09-101-272G-62
Sequence 62, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 62
LENGTH: 89
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
US-09-101-272G-62

Query Match 64.7%; Score 513; DB 4; Length 89;
Best Local Similarity 100.0%; Pred. No. 2e-44;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 BIDSKTCYEGNGHYFGKASTDTMGPRCLPWNATVLQOTYHAHRSDALQLGLGKHNYC 102
DB 1 BIDSKTCYEGNGHYFGKASTDTMGPRCLPWNATVLQOTYHAHRSDALQLGLGKHNYC 60
QY 103 RNPONRRPWCYVQVGLKPLVQECWVHDC 131
|||||

DB 61 RNPONRRPWCYVQVGLKPLVQECWVHDC 89
RESULT 23
US-08-560-098A-51
Sequence 51, Application US/08560098A
Patent No. 5976841
GENERAL INFORMATION:
APPLICANT: WENDEL, Stephan
APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 477 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-560-098A-51

Query Match 42.3%; Score 335.5; DB 2; Length 477;
Best Local Similarity 46.3%; Pred. No. 8.3e-26;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;
QY 3 ELHQVP---SNCDCLNGTGVSNKYFSNIHWCNCPKFGGHCHEIDKSKTCYEGNGHYF 58
DB 78 QCHTVFKVSCSELRCFNGGTQWQAASFSDF-VCCPKGYTGKQCEVDTHATCYKQGVY 136

QY 59 RKGASTDTMGPRCLPWNATVLQOTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVG 118
DB 137 RGTWSTSESQAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPNNSKPCYVIKA 196
QY 119 LKPLVQECWVHDC 132
DB 197 SKFILEFCVPVCS 210

RESULT 24
US-07-609-510B-16
Sequence 16, Application US/07609510B
Patent No. 5326700
GENERAL INFORMATION:
APPLICANT: Berg et al.
TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue I
NUMBER OF SEQUENCES: 16

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: IN
COUNTRY: U.S.A.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,510B
FILING DATE: 19901106
CLASSIFICATION: 435
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 41.4%; Score 328.5; DB 1; Length 527;
Best Local Similarity 46.3%; Pred. No. 4.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELQVPS-SNCD---CLNGGTCVSNKYFNSIHWNCNPKFGGQHCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVFVKSCSEPCFNGGTCQALYFSDP-VQCPEGFAGKCCBEIDTRATCYEDQGYSY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPNRRRRCWCVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAKPKYSGRRPDAILRLGLGNHNYCRNPDNRDSDKPCWCVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

```

```

Query Match 41.4%; Score 328.5; DB 1; Length 527;
Best Local Similarity 46.3%; Pred. No. 4.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELQVPS-SNCD---CLNGGTCVSNKYFNSIHWNCNPKFGGQHCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVFVKSCSEPCFNGGTCQALYFSDP-VQCPEGFAGKCCBEIDTRATCYEDQGYSY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPNRRRRCWCVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAKPKYSGRRPDAILRLGLGNHNYCRNPDNRDSDKPCWCVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

```

```

RESULT 25
US-08-811-949-39
Sequence 39, Application US/08811949
Patent No. 5840533
GENERAL INFORMATION:
APPLICANT: NIWA, MINEO
APPLICANT: SAITO, YOSHIMASA
APPLICANT: SASAKI, HITOSHI
APPLICANT: HAYASHI, MASAKO
APPLICANT: NOTANI, JOUJI
APPLICANT: KOBAYASHI, MASAKAZU
TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESS: P.C. JEFFERSON DAVIS HIGHWAY, SUITE 400
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/811,949
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

```

```

NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 18-966-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-811-949-39

Query Match 41.4%; Score 328.5; DB 2; Length 527;
Best Local Similarity 46.3%; Pred. No. 4.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELQVPS-SNCD---CLNGGTCVSNKYFNSIHWNCNPKFGGQHCEIDKSKTCYEGNGHFY 58
Db 42 QCHSVFVKSCSEPCFNGGTCQALYFSDP-VQCPEGFAGKCCBEIDTRATCYEDQGYSY 100
QY 59 RGKASTDTMGRPCLPWNSATVLQOTYHAHRSDALQGLGKHNCRPNRRRRCWCVQVG 118
Db 101 RGTWSTAESGAECTNWNSSALAKPKYSGRRPDAILRLGLGNHNYCRNPDNRDSDKPCWCVFKA 160
QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

```

```

RESULT 26
PCT-US91-01025A-2
Sequence 2, Application PC/TUS9101025A
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
TITLE OF INVENTION: Tissue plasminogen Activator Having Fibrin
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/01025A
FILING DATE: 19910214
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/486,657
FILING DATE: 1 March 1990
ATTORNEY/AGENT INFORMATION:
NAME: Hasak, Janet B.
REGISTRATION NUMBER: 28,616
REFERENCE/DOCKET NUMBER: 454P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/266-1896
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 527 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
PCT-US91-01025A-2

```



```

; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-966-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 562 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-811-949-43

Query Match 41.4%; Score 328.5; DB 2; Length 562;
Best Local Similarity 46.3%; Pred. No. 5.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYESNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQQALYFSDP-VQCPEGFAGKCCIEDTRATCYEDQGISY 135

QY 59 RGRKASDTMGRCPLPWN SATVLQQT YHAHRSDALQLGLGKHN YCRNPDRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNNSSALAKPYSGRPPDAIRLGLGNHNYCRNPDRDRDSKPCYVFKA 195

QY 119 LKPLVOECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211

```

Search completed: December 3, 2003, 14:45:35
Job time : 7.98276 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 21.2657 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-4
Perfect score: 793
Sequence: 1 SNELHQVPNCCLNGGTCV.....QVGLKPLVQECMVHDCADGK 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	135	AAE16545	Human urokinase-ty
2	793	100.0	143	AAE16549	Human uPA amino te
3	793	100.0	337	AAE16549	Human colon cancer
4	793	100.0	337	AAE16549	Human ovarian anti
5	793	100.0	403	AAE16547	Human urokinase-ty
6	793	100.0	411	AAE16547	Sequence encoded b
7	793	100.0	411	AAE16547	Human native prour
8	793	100.0	411	AAE16547	UK-S3 as encoded b
9	793	100.0	411	AAE16547	Urokinase precursor

10	793	100.0	411	AAE10057	Pro-urokinase deri
11	793	100.0	411	AAE10058	Pro-urokinase deri
12	793	100.0	411	AAE10058	PUK. Homo sapiens
13	793	100.0	411	AAE10058	Pro-urokinase. Ho
14	793	100.0	411	AAE10058	Pro-urokinase muta
15	793	100.0	411	AAE10058	Pro-urokinase muta
16	793	100.0	411	AAE10058	Pro-urokinase muta
17	793	100.0	411	AAE10058	Pro-urokinase muta
18	793	100.0	411	AAE10058	Pro-urokinase muta
19	793	100.0	411	AAE10058	Pro-urokinase muta
20	793	100.0	411	AAE10058	Pro-urokinase muta
21	793	100.0	411	AAE10058	Pro-urokinase muta
22	793	100.0	411	AAE10058	Pro-urokinase muta
23	793	100.0	411	AAE10058	Pro-urokinase muta
24	793	100.0	411	AAE10058	Pro-urokinase muta
25	793	100.0	411	AAE10058	Pro-urokinase muta
26	793	100.0	411	AAE10058	Pro-urokinase muta
27	793	100.0	411	AAE10058	Pro-urokinase muta
28	793	100.0	411	AAE10058	Pro-urokinase muta
29	793	100.0	411	AAE10058	Pro-urokinase muta
30	793	100.0	411	AAE10058	Pro-urokinase muta
31	793	100.0	411	AAE10058	Pro-urokinase muta
32	793	100.0	411	AAE10058	Pro-urokinase muta
33	793	100.0	411	AAE10058	Pro-urokinase muta
34	793	100.0	411	AAE10058	Pro-urokinase muta
35	793	100.0	411	AAE10058	Pro-urokinase muta
36	793	100.0	411	AAE10058	Pro-urokinase muta
37	793	100.0	411	AAE10058	Pro-urokinase muta
38	793	100.0	411	AAE10058	Pro-urokinase muta
39	793	100.0	411	AAE10058	Pro-urokinase muta
40	793	100.0	411	AAE10058	Pro-urokinase muta
41	793	100.0	411	AAE10058	Pro-urokinase muta
42	793	100.0	411	AAE10058	Pro-urokinase muta
43	793	100.0	411	AAE10058	Pro-urokinase muta
44	793	100.0	411	AAE10058	Pro-urokinase muta
45	793	100.0	411	AAE10058	Pro-urokinase muta

ALIGNMENTS

RESULT 1
AAE16545
ID AAE16545 standard; Protein; 135 AA.
AC AC
AAE16545;
XX AC
XX 09-APR-2002 (first entry)
XX XX
XX Human urokinase-type plasminogen activator amino terminal fragment (ATF).
XX KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
XX KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
XX KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX KW clotting disorder; uterine contraction disorder; respiratory disease;
XX KW adult respiratory distress syndrome; amino terminal fragment; ATF;
XX male impotence.
XX OS
XX Homo sapiens.
XX XX
XX WO200197752-A2.
XX XX
XX 27-DEC-2001.
XX XX
XX 13-JUN-2001; 2001WO-US18976.
XX XX
XX 20-JUN-2000; 2000US-212874P.
XX XX
XX (UYPE-) UNIV PENNSYLVANIA.
XX PA
XX Human native prour
XX UK-S3 as encoded b
XX PI
XX Cines DB, Higazi AA;

DR WPI; 2002-122240/16.
DR N-PSDB; AAD27078.
XX
PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
XX Claim 11; Fig 1D; 117pp; English.
PS
XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) amino
CC terminal fragment (ATF).
XX
SQ Sequence 135 AA;
Query Match 100.0%; Score 793; DB 23; Length 135;
Best Local Similarity 100.0%; Pred. No. 7.4e-52;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQPVSNCDCCLNGTGVSNKYFNSIHWNCNCPKFGGHCIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPVSNCDCCLNGTGVSNKYFNSIHWNCNCPKFGGHCIDKSKTCYEGNGHFYRG 60
Qy 61 KASDTMGRPCLPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 61 KASDTMGRPCLPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Qy 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135
RESULT 2
AAE16549
ID AAE16549 standard; Protein; 143 AA.
AC AAE16549;
DT 09-APR-2002 (first entry)
XX
XX Human uPA amino terminal fragment (ATF) and connecting peptide.
DE
XX Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
XX microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
XX tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
XX clotting disorder; uterine contraction disorder; respiratory disease;
XX adult respiratory distress syndrome; amino terminal fragment; ATF;
XX male impotence.
XX
XX Homo sapiens.
XX
XX WO2000197752-A2.
XX
XX 27-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-US18976.
XX
XX 20-JUN-2000; 2000US-212874P.
XX
XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;
XX
XX WPI; 2002-122240/16.
DR N-PSDB; AAD27082.
XX
PT Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX
XX Claim 24; Fig 1H; 117pp; English.
PS
XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) amino
CC terminal fragment (ATF) and connecting peptide.
XX
SQ Sequence 143 AA;
Query Match 100.0%; Score 793; DB 23; Length 143;
Best Local Similarity 100.0%; Pred. No. 7.8e-52;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQPVSNCDCCLNGTGVSNKYFNSIHWNCNCPKFGGHCIDKSKTCYEGNGHFYRG 60
Db 1 SNELHQPVSNCDCCLNGTGVSNKYFNSIHWNCNCPKFGGHCIDKSKTCYEGNGHFYRG 60
Qy 61 KASDTMGRPCLPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 61 KASDTMGRPCLPWNSATVLQOTVHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Qy 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135
RESULT 3
AAG75492
ID AAG75492 standard; Protein; 337 AA.
AC AAG75492;
DT 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen protein SEQ ID NO:6256.
DE
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; chromosome 10.
XX
XX Homo sapiens.
XX
XX WO2000122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US26524.
XX
XX 29-SEP-1999; 99US-0157137.
PR 03-NOV-1999; 99US-0163280.
XX
XX (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;
 XX WPI; 2001-225357/24.
 DR N-PSDB; AAH34897.
 XX
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 XX Claim 11; Page 7707-7708; 9803pp; English.
 PS
 XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAH77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SQ Sequence 337 AA;
 Query Match 100.0%; Score 793; DB 22; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1.6e-51;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 60
 DB 27 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 86
 QY 61 KASTDTMGRCPLPMSATVLQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 87 KASTDTMGRCPLPMSATVLQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 146
 QY 121 PLVQECMVHDCADGK 135
 DB 147 PLVQECMVHDCADGK 161
 RESULT 4
 ID ABP41795 standard; protein; 337 AA.
 XX
 AC ABP41795;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HWCEB79, SEQ ID NO:2927.
 XX
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KW inflammatory condition; immune disorder; blood disorder;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;
 KW gastrointestinal disorder; urinary system disorder; drug screening;
 KW gene therapy; chromosome mapping; forensic analysis;
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KW antiinflammatory; gynaecological; reproductive; chromosome 10q24.
 XX
 OS Homo sapiens.
 XX
 XX WO200200677-A1.
 PN

XX 03-JAN-2002.
 XX
 XX 07-JUN-2001; 2001WO-US18569.
 XX
 XX 07-JUN-2000; 2000US-209467P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Birse CE, Rosen CA;
 XX
 XX WPI; 2002-147878/19.
 DR N-PSDB; ABQ54872.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 XX Claim 11; SEQ ID NO 2927; 2922pp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIFO
 CC at ftp.wifo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 337 AA;
 Query Match 100.0%; Score 793; DB 23; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1.6e-51;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 60
 DB 27 SNEHQVPSNCDCLNGTGVSNKYFNSHWCNCPKPKFGGHCIDKSKTCYEGNGHFYRG 86
 QY 61 KASTDTMGRCPLPMSATVLQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 120
 DB 87 KASTDTMGRCPLPMSATVLQOTVHAHRSALQGLGKHNYCRNPNRRPWCYVQVGLK 146
 QY 121 PLVQECMVHDCADGK 135
 DB 147 PLVQECMVHDCADGK 161
 RESULT 5
 AAEL16547
 ID AAEL16547 standard; protein; 403 AA.

XX AAE16547;
AC 09-APR-2002 (first entry)
XX Human urokinase-type plasminogen activator scUPA delta136-143 mutant.
XX
DE Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
XX stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
KW clotting disorder; uterine contraction disorder; respiratory disease;
KW male impotence; adult respiratory distress syndrome; scUPA delta136-143;
XX single chain urokinase; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX WO200197752-A2.
XX 27-DEC-2001.
XX 13-JUN-2001; 2001WO-US18976.
XX 20-JUN-2000; 2000US-212874P.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Cines DB, Higazi AA;
XX WPI; 2002-122240/16.
XX N-PSDB; AAD27080.
XX Composition for modulating muscle cell and tissue contractility for
PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
PT comprising domains from urokinase-type plasminogen activator -
XX Claim 22; Fig 1F; 117pp; English.
XX The invention relates to a composition comprising one or more domains of
CC urokinase-type plasminogen activator (uPA). The composition is used to
CC modulate the contractility and angiogenic activity of a mammalian muscle,
CC endothelial cell or tissue. The composition is used for treating stroke,
CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
CC diabetic retinopathy, wound healing, clotting disorder, uterine
CC contraction disorder, male impotence, respiratory disease or condition
CC such as asthma, adult respiratory distress syndrome, primary pulmonary
CC hypertension, microvascular thrombotic occlusion, and a disorder
CC associated with chronic intrapulmonary fibrin formation. The present
CC sequence is human urokinase-type plasminogen activator (uPA) single
CC chain urokinase (scUPA) deletion mutant designated as scUPA delta136-143.
XX
SQ Sequence 403 AA;
Query Match 100.0%; Score 793; DB 23; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHCNCPKFGGQHCIDSKTCYEGNGHGYRG 60
DB 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHCNCPKFGGQHCIDSKTCYEGNGHGYRG 60
QY 61 KASTDTWGRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPDRRPPWCYVOGLK 120
DB 61 KASTDTWGRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPDRRPPWCYVOGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 6
AAP50871
ID AAP50871 standard; protein; 411 AA.
XX
XX AAP50871;
XX 30-NOV-1991 (first entry)
DT Sequence encoded by cDNA sequence for human urokinase zymogen
XX (Japanese Patent Application No.37119/84).
DE Thrombolytic agent; plasminogen activator activity; fibrin affinity;
KW enzyme.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Cleavage-site 158..159
FT /note= "potential cleavage site which generates
FT the two-chain form from the zymogen"
FT
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 233..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
XX EP139447-A.
XX
XX 02-MAY-1985.
XX
XX 07-SEP-1984; 84EP-0306117.
XX
XX 17-OCT-1983; 83JP-0195051.
XX 13-SEP-1983; 83JP-0170354.
XX
XX (GREG) GREEN CROSS CORP.
XX
XX Kasai S, Arimura H, Mori K, Nishida M, Suyama T;
XX WPI; 1985-106530/18.
XX New urokinase zymogen - useful as thrombolytic agent
XX
XX Disclosure; Page 12; 30pp; English.
XX Zymogen AAP50871 is the inactive precursor form of human urokinase.
XX Urokinase zymogen is cleaved into the two-chain form composed of
XX characteristic urokinase H (molecular wt. of 30,000) and L (molecular
XX wt. of 20,000) chains when treated with catalytic amounts of plasmin.
XX The patentors claim a new urokinase zymogen which has mol. wt. ca.
XX 50,000, a single chain molecular structure, and selective affinity
XX for fibrin. It is a thrombolytic agent which manifests its
XX plasminogen activator activity on cleavage by proteolytic enzymes
XX (e.g. plasmin) and has higher affinity for fibrin than known forms
XX of urokinase.
XX
SQ Sequence 411 AA;
Query Match 100.0%; Score 793; DB 6; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHCNCPKFGGQHCIDSKTCYEGNGHGYRG 60
DB 1 SNELHQVPSNCDCLNGTCTVSNKYFNSIHCNCPKFGGQHCIDSKTCYEGNGHGYRG 60
QY 61 KASTDTWGRCLPWNATVLOQTYHAHRSDALQLGLGKHYCRNPDRRPPWCYVOGLK 120

Db 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGK 135

Db 121 PLVQECMVHDCADGK 135

RESULT 7

AAW13634

ID AAW13634 standard; Protein; 411 AA.

AC AAW13634;

XX 04-JUN-1997 (first entry)

XX Human native prourokinase.

XX Human; prourokinase; hPUK; variant; half-life; increase; EGF;

XX epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Domain 10..42

XX /label= EGF domain

XX /note= "in_Glained variants, at least part of the

XX EGF domain is deleted (see comments)"

XX Region 10..19

XX /label= first_loop

XX Region 20..31

XX /label= second_loop

XX Region 33..42

XX /label= third_loop

XX EP398361-A.

XX 22-NOV-1990.

XX 18-MAY-1990; 90EP-0109472.

XX 22-FEB-1990; 90JP-0042020.

XX 18-MAY-1989; 89JP-0126433.

XX 03-JUL-1986; 86JP-0156936.

XX 18-FEB-1987; 87JP-0036495.

XX 18-MAY-1989; 89JP-0126434.

XX (GREG) GREEN CROSS CORP.

XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;

XX Morita M, Tanabe T;

XX WPI; 1990-350146/47.

XX N-PSDB; AAT61671.

XX Human pro-urokinase variants - deficient in loop regions of

XX epidermal growth factor, showing long blood half-life, as

XX fibrinolytic agent

XX Disclosure; Fig 1; 22pp; English.

XX New variants of human prourokinase (hPUK) comprise a hPUK deficient

XX in (i) at least part of the first loop region of the epidermal growth

XX factor (EGF) domain; (ii) at least part of the first loop and at

XX least part of the second loop; or (iii) at least part of the third

XX loop. The hPUK variants show an increased blood half-life comparable

XX to that of the whole EGF domain-deficient hPUK variant and urokinase

XX while retaining the same properties as those of hPUK. They have

XX potent thrombolytic activity and very little tendency to cause

XX spontaneous bleeding. The present sequence is that of the wild-type

XX hPUK protein, including the EGF domain.

XX Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;

Best Local Similarity 100.0%; Pred. No. 1.9e-51;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

Db 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

QY 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120

Db 61 KASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120

QY 121 PLVQECMVHDCADGK 135

Db 121 PLVQECMVHDCADGK 135

RESULT 8

AAW05117

ID AAR05117 standard; protein; 411 AA.

XX AAR05117;

XX 25-MAR-2003 (updated)

XX 04-OCT-1990 (first entry)

XX UK-S3 as encoded by PUKS3.

XX Urokinase; glycosylation.

XX Homo sapiens.

XX Key Location/Qualifiers

XX misc_difference 153

XX /label=synthetic_mutation

XX /note="old seq (Ieu)"

XX misc_difference 155

XX /label=synthetic_mutation

XX /note="old seq (Pro)"

XX EP370205-A.

XX 30-MAY-1990.

XX 28-SEP-1989; 89EP-0117981.

XX 29-SEP-1988; 88JP-0245705.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Sasaki K, Nishi T, Yasumura S, Sato M, Itoh S;

XX WPI; 1990-165029/22.

XX N-PSDB; AAQ04486.

XX Polypeptide(s) with added carbohydrate chains - formed by

XX modification of amino acid sequence, used to improve

XX physicochemical properties and/or activities.

XX Disclosure; Page ?; 30pp; English.

XX The polypeptide is a deriv. of mature urokinase, designated

XX UK-S3 which has 2 amino acid substns. which result in an N-linked

XX glycosylation site giving the new protein improved stability and

XX activity.

XX See also AAR05113-17.

XX (Updated on 25-MAR-2003 to correct PR field.)

XX (Updated on 25-MAR-2003 to correct PA field.)

XX (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;

Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 9
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX AAR06244;
AC AAR06244;
XX 07-DEC-1990 (first entry)
DT Urokinase precursor protein.
DE Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
XX KW myocardial infarction;
KW myocardi al infarction;
XX OS Homo sapiens.
XX PI EP380334-A.
PN 01-AUG-1990.
PD 25-JAN-1990; 90EP-0300772.
PF 17-MAY-1989; 89JP-0121405.
PR 27-JAN-1989; 89JP-0016406.
XX (GREC) GREEN CROSS CORP.
XX PI Matsuda H, Ueda Y, Tamanouchi K;
XX WPI; 1990-233117/31.
XX Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bio:availability and improved activity
XX
PS Claim 3; Fig 1; 1lpp; English.
XX By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135

RESULT 9
AAR06244
ID AAR06244 standard; protein; 411 AA.
XX AAR06244;
AC AAR06244;
XX 07-DEC-1990 (first entry)
DT Urokinase precursor protein.
DE Urokinase precursor; fibrinolysis; thrombolytic; cerebral thrombosis;
XX KW myocardial infarction;
KW myocardi al infarction;
XX OS Homo sapiens.
XX PI EP380334-A.
PN 01-AUG-1990.
PD 25-JAN-1990; 90EP-0300772.
PF 17-MAY-1989; 89JP-0121405.
PR 27-JAN-1989; 89JP-0016406.
XX (GREC) GREEN CROSS CORP.
XX PI Matsuda H, Ueda Y, Tamanouchi K;
XX WPI; 1990-233117/31.
XX Urokinase precursor-lipid composite - used as thrombolytic agent,
PT having prolonged half-life in the blood, enhanced
PT bio:availability and improved activity
XX
PS Claim 3; Fig 1; 1lpp; English.
XX By forming a precursor-lipid composite, the half-life of this
CC thrombolytic agent in the blood may be increased, exhibiting
CC improved activity without abnormal acceleration of fibrinolytic
CC activity. Compound is useful as a thrombolytic agent in
CC treatment of cerebral thrombosis, myocardial infarction etc.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135

DB 121 PLVQECMVHDCADGK 135

RESULT 10
AAR10057
ID AAR10057 standard; Protein; 411 AA.
XX AAR10057;
AC AAR10057;
XX 18-MAR-1991 (first entry)
DT Pro-uokinase derivative UK-T4 with Ser(153) and Thr(155).
DE pro-uokinase; UK-T4; plasminogen activator; myocardial infarction;
XX KW cerebral thrombosis.
KW cerebral thrombosis.
XX OS Homo sapiens.
XX PN EP405285-A.
XX 02-JAN-1991.
XX 18-JUN-1990; 90EP-0111471.
XX 19-JUN-1989; 89JP-0156302.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Yasamura S, Nishi T, Ito S;
XX WPI; 1991-008678/02.
XX N-PSDB; AAQ10169.
XX New plasminogen activator almost identical to natural
PT pro-uokinase - is thrombin resistant and used for
PT prophylaxis, treatment of cerebral thrombosis or myocardial
XX infarction
XX Disclosure; Page 8; 84pp; English.
XX UK-T4 is one example of a plasminogen activator which differs from
CC natural human pro-uokinase at positions 153 and 155 (Leu
CC substituted by Ser; Pro substituted by Thr, respectively).
CC The derivative has decreased susceptibility to thrombin compared to
CC natural type pro-UK and higher specific activity.
CC See also AAQ10168 and AAQ10170.
XX
SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 12; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
DB 61 KASTDTMGRPCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 11
AAR10058
ID AAR10058 standard; Protein; 411 AA.
XX AAR10058;
AC AAR10058;
XX

DT 18-MAR-1991 (first entry)
 XX Pro-prokinase derivative UK-S3 with Asn(153) and Thr(155).
 XX pro-urokinase; UK-S3; plasminogen activator; myocardial infarction;
 KW cerebral thrombosis.
 XX Homo sapiens.
 XX EP405285-A.
 XX 02-JAN-1991.
 PD 18-JUN-1990; 90EP-0111471.
 XX 19-JUN-1989; 89JP-0156302.
 PR (KYOW) KYOWA HAKKO KOGYO KK.
 XX Yasamura S, Nishi T, Ito S;
 PI WPI; 1991-008678/02.
 DR N-PSDB; AAQ10170.
 DR New plasminogen activator almost identical to natural
 PT pro-urokinase - is thrombin resistant and used for
 PT prophylaxis, treatment of cerebral thrombosis or myocardial
 PT infarction
 XX Disclosure; Page 9; 84pp; English.
 XX UK-S3 is one example of a plasminogen activator which differs from
 CC natural human pro-urokinase at positions 153 and 155. (Leu
 CC substituted by Asn; Pro substituted by Thr, respectively)
 CC The derivative has decreased susceptibility to thrombin compared to
 CC natural type pro-UK and higher specific activity.
 CC See also AAQ10168 and AAQ10169.
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 793; DB 12; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.9e-51;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
 QY 121 PLVQECMVHDCADGK 135
 Db 121 PLVQECMVHDCADGK 135
 RESULT 12
 AAR40225
 ID AAR40225 standard; Protein; 411 AA.
 AC AAR40225;
 XX 10-FEB-1994 (first entry)
 DT PUK.
 DE Pre-urokinase; thrombolytic; blood; plasmid; PUK.
 KW Homo sapiens.
 XX Key Location/Qualifiers
 XX Region 10..60
 FT

FT /note= "Ser, Asn, Pro, Gly, or Tyr in the region
 FT 10 to 60 in the N-terminal of the human PUK
 FT can be replaced by Thr, Pro or Ala"
 FT Misc-difference 64
 FT /note= "Thr encoded by AGT (sic)"
 FT Misc-difference 177
 FT /note= "Thr encoded by TAC (sic)"
 FT Domain 10..49
 FT /label= EGF
 FT Region 10..19
 FT /label= Loop_1
 FT Region 20..31
 FT /label= Loop_2
 FT Region 33..42
 FT /label= Loop_3
 XX JP05192142-A.
 PN 03-AUG-1993.
 XX 20-JAN-1992; 92JP-0030178.
 PF 20-JAN-1992; 92JP-0030178.
 XX (GREC) GREEN CROSS CORP.
 PA WPI; 1993-277461/35.
 XX N-PSDB; AAQ48228.
 DR Mutant human pre-urokinase - by replacing specified aminoacid(s)
 XX in N-terminal for providing longer half-life in blood and higher
 PT thrombolytic ability
 PT Claim 1; Page 14-16; 26pp; Japanese.
 PS Human pre-urokinase can be modified by replacing: Ser, Asn,
 CC Pro, Gly, or Tyr in the region ranging from residue 10 to 60 in the
 CC N-terminal of the human PUK by Thr, Pro or Ala. These mutants have a
 CC longer half-life in the blood and a higher thrombolytic ability.
 CC For examples see (AAR47956-R47960).
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 793; DB 14; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.9e-51;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
 Db 1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGQHCIEDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
 Db 61 KASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPNRRRPMCYVQVGLK 120
 QY 121 PLVQECMVHDCADGK 135
 Db 121 PLVQECMVHDCADGK 135
 RESULT 13
 AAR62991
 ID AAR62991 standard; protein; 411 AA.
 XX AAR62991;
 XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 DE Pro-urokinase.
 XX Pro-urokinase; thrombolysis; fibrin clot lysis.
 XX

```

OS Homo sapiens.
XX Key Location/Qualifiers
FH Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07278.
XX 02-JUL-1993; 93US-0087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX Disclosure; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
XX Query Match 100.0%; Score 793; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 1 SNELHQPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPNRRPWCYVOVGLK 120
Db 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPNRRPWCYVOVGLK 120
QY 121 PLVQECWVHDCADGK 135
Db 121 PLVQECWVHDCADGK 135
RESULT 14
AAR62992
ID AAR62992 standard; protein; 411 AA.
XX AAR62992;
XX AC
XX XX
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)

```

```

XX DE Pro-urokinase mutant Ala300.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala300;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX OS Homo sapiens.
XX Key Location/Qualifiers
FH Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07278.
XX 02-JUL-1993; 93US-0087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
XX fibrinogenolysis activity and non-specific plasminogen activation
XX Claim 5; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
XX described in AAR62992-R63008 were derived. These mutants retain the
XX thrombolytic activity of the wild type protein, useful for the
XX treatment of thromboembolism, but have a reduced fibrinogenolysis
XX activity and non-specific plasminogen activation. The mutants can
XX therefore be used for the lysis of fibrin clots without inducing
XX systemic bleeding, as can be the case with the wild type protein.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
XX Query Match 100.0%; Score 793; DB 16; Length 411;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-51;
XX Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
Db 1 SNELHQPNSCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPNRRPWCYVOVGLK 120
Db 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRPNPNRRPWCYVOVGLK 120
QY 121 PLVQECWVHDCADGK 135
Db 121 PLVQECWVHDCADGK 135
RESULT 15

```

AAR62993
ID AAR62993 standard; protein; 411 AA.
XX AC AAR62993;
XX 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant His300.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His300;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07278.
XX 02-JUL-1993; 93US-0087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation
XX Claim 5; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R6308 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDLNGGTCVSNKYFSNTHWCNPKKFGQHCETDKSKTCYEGNGHYRG 60
DB 1 SNELHQVPSNCDLNGGTCVSNKYFSNTHWCNPKKFGQHCETDKSKTCYEGNGHYRG 60
QY 61 KASTDTMGRPCLPWNASVLTQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQGLK 120
DB 61 KASTDTMGRPCLPWNASVLTQTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQGLK 120

QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135
RESULT 16
AAR62994
ID AAR62994 standard; protein; 411 AA.
XX AC AAR62994;
XX 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX DE Pro-urokinase mutant His299 Ala300.
XX KW Pro-urokinase; thrombolysis; fibrin clot lysis;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding; mutant His 299 Ala300.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 297..313
FT Disulfide-bond /note= "flexible loop"
FT 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX WO9501427-A1.
XX 12-JAN-1995.
XX 28-JUN-1994; 94WO-US07278.
XX 02-JUL-1993; 93US-0087163.
XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX Gurewich V, Liu J;
XX WPI; 1995-060991/08.
XX Pro-urokinase mutants - have thrombolytic activity but reduced
PT fibrinogenolysis activity and non-specific plasminogen activation
XX Claim 7; Fig 1; 46pp; English.
XX AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R6308 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 411 AA;
Query Match 100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPDCCLNGGTCVSNKYFSNIHWNCPCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 SNELHQPSPDCCLNGGTCVSNKYFSNIHWNCPCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 KASDTMTGRPCLPWNSATVLIQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 KASDTMTGRPCLPWNSATVLIQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 PLVQECMVHDCADGK 135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 PLVQECMVHDCADGK 135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 17

AAR62995
 ID AAR62995 standard; protein; 411 AA.

XX AC AAR62995;

XX AC AAR62995;

DT 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant Ala300 Ala301.

DE Pro-urokinase; thrombolysis; fibrin clot lysis;
 XX reduced fibrinogenolysis; non-specific plasminogen activation;
 KW mutant Ala300 Ala301; systemic bleeding.
 KW
 KW
 OS Homo sapiens.

XX Key

PH Domain

297..313 Location/Qualifiers

/note= "flexible loop"

FT Disulfide-bond 11..19

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced

XX fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 9; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

XX described in AAR62992-R63008 were derived. These mutants retain the

XX thrombolytic activity of the wild type protein, useful for the

XX treatment of thromboembolism, but have a reduced fibrinogenolysis

XX activity and non-specific plasminogen activation. The mutants can

XX therefore be used for the lysis of fibrin clots without inducing

XX systemic bleeding, as can be the case with the wild type protein.

XX (Updated on 25-MAR-2003 to correct PN field.)

XX SQ

Sequence 411 AA;

Query Match 100.0%; Score 793; DB 16; Length 411;

Best Local Similarity 100.0%; Pred. No 1.9e-51;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQPSPDCCLNGGTCVSNKYFSNIHWNCPCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 SNELHQPSPDCCLNGGTCVSNKYFSNIHWNCPCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 61 KASDTMTGRPCLPWNSATVLIQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 KASDTMTGRPCLPWNSATVLIQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVGLK 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 121 PLVQECMVHDCADGK 135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 121 PLVQECMVHDCADGK 135
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 18

AAR62996

ID AAR62996 standard; protein; 411 AA.

XX AC AAR62996;

XX AC AAR62996;

DT 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX Pro-urokinase mutant His300 Ala301.

DE Pro-urokinase; thrombolysis; fibrin clot lysis;
 XX reduced fibrinogenolysis; non-specific plasminogen activation;
 KW mutant His300 Ala301; systemic bleeding.
 KW
 KW
 OS Homo sapiens.

XX Key

PH Domain

297..313 Location/Qualifiers

/note= "flexible loop"

FT Disulfide-bond 11..19

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

XX WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced

XX fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 9; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

XX described in AAR62992-R63008 were derived. These mutants retain the

XX thrombolytic activity of the wild type protein, useful for the

XX treatment of thromboembolism, but have a reduced fibrinogenolysis

XX activity and non-specific plasminogen activation. The mutants can

XX therefore be used for the lysis of fibrin clots without inducing

XX systemic bleeding, as can be the case with the wild type protein.

XX (Updated on 25-MAR-2003 to correct PN field.)

CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
SQ

Query Match 100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVOVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVOVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 19
AAR62997
ID AAR62997 standard; protein; 411 AA.
XX
AC AAR62997;
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Gly306.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Gly306;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 297..313 /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
PN WO9501427-A1.
XX
PD 12-JAN-1995.
XX
PF 28-JUN-1994; 94WO-US07278.
XX
PR 02-JUL-1993; 93US-0087163.
XX
PA (NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
XX
PI Gurewich V, Liu J;
XX
DR WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced
FT fibrinogenolysis activity and non-specific plasminogen activation
XX
PS Claim 13; Fig 1; 46pp; English.
XX
CC AAR62991 is the wild type pro-urokinase, from which the new mutants
CC described in AAR62992-R63008 were derived. These mutants retain the
CC thrombolytic activity of the wild type protein, useful for the
CC treatment of thromboembolism, but have a reduced fibrinogenolysis
CC activity and non-specific plasminogen activation. The mutants can
CC therefore be used for the lysis of fibrin clots without inducing
CC systemic bleeding, as can be the case with the wild type protein.
CC (updated on 25-MAR-2003 to correct PN field.)
XX Sequence 411 AA;
SQ

Query Match 100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVOVGLK 120
DB 61 KASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPMCYVOVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 20
AAR62998
ID AAR62998 standard; protein; 411 AA.
XX
AC AAR62998;
DT 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ala313.
XX
KW Pro-urokinase; thrombolysis; fibrin clot lysis; mutant Ala313;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 297..313 /note= "flexible loop"
FT Disulfide-bond 11..19
FT Disulfide-bond 13..31
FT Disulfide-bond 33..42
FT Disulfide-bond 50..131
FT Disulfide-bond 71..113
FT Disulfide-bond 102..126
FT Disulfide-bond 148..279
FT Disulfide-bond 189..205
FT Disulfide-bond 197..268
FT Disulfide-bond 293..362
FT Disulfide-bond 325..341
FT Disulfide-bond 352..380
XX
PN WO9501427-A1.
XX
PD 12-JAN-1995.
XX
PF 28-JUN-1994; 94WO-US07278.
XX

PR 02-JUL-1993; 93US-0087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Gurewich V, Liu J;
 PI WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis
 CC activity and non-specific plasminogen activation. The mutants can
 CC therefore be used for the lysis of fibrin clots without inducing
 CC systemic bleeding, as can be the case with the wild type protein.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 793; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.9e-51; Indels 0; Gaps 0;
 Matches 135; Conservative 0; Mismatches 0;
 QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
 DB 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRPRWCYVQVGLK 120
 DB 61 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRPRWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGK 135
 DB 121 PLVQECMVHDCADGK 135
 RESULT 21
 AAR62999
 ID AAR62999 standard; protein; 411 AA.
 XX AAR62999;
 XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant His313.
 DE Pro-urokinase; thrombolysis; fibrin clot lysis; mutant His313;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT Domain 297..313 /note= "flexible loop"
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113
 FT Disulfide-bond 102..126
 FT Disulfide-bond 148..279
 FT Disulfide-bond 189..205
 FT Disulfide-bond 197..268
 FT Disulfide-bond 293..362
 FT Disulfide-bond 325..341
 FT Disulfide-bond 352..380

XX WO9501427-A1.
 XX 12-JAN-1995.
 PD 28-JUN-1994; 94WO-US07278.
 XX 02-JUL-1993; 93US-0087163.
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XX Gurewich V, Liu J;
 PI WPI; 1995-060991/08.
 XX Pro-urokinase mutants - have thrombolytic activity but reduced
 PT fibrinogenolysis activity and non-specific plasminogen activation
 XX Claim 11; Fig 1; 46pp; English.
 XX AAR62991 is the wild type pro-urokinase, from which the new mutants
 CC described in AAR62992-R63008 were derived. These mutants retain the
 CC thrombolytic activity of the wild type protein, useful for the
 CC treatment of thromboembolism, but have a reduced fibrinogenolysis
 CC activity and non-specific plasminogen activation. The mutants can
 CC therefore be used for the lysis of fibrin clots without inducing
 CC systemic bleeding, as can be the case with the wild type protein.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 411 AA;
 SQ Query Match 100.0%; Score 793; DB 16; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.9e-51; Indels 0; Gaps 0;
 Matches 135; Conservative 0; Mismatches 0;
 QY 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
 DB 1 SNEHQVPSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFYRG 60
 QY 61 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRPRWCYVQVGLK 120
 DB 61 KASTDTMGPRCLPWN SATVLOQTYHAHRS DALQLGLGKHNYCRNPDRRPRWCYVQVGLK 120
 QY 121 PLVQECMVHDCADGK 135
 DB 121 PLVQECMVHDCADGK 135
 RESULT 22
 AAR63000
 ID AAR63000 standard; protein; 411 AA.
 XX AAR63000;
 XX 25-MAR-2003 (updated)
 DT 21-SEP-1995 (first entry)
 XX Pro-urokinase mutant Ser175 His187.
 DE Pro-urokinase; thrombolysis; fibrin clot lysis;
 KW reduced fibrinogenolysis; non-specific plasminogen activation;
 KW systemic bleeding; mutant Ser175 His187.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT Domain 297..313 /note= "flexible loop"
 FT Disulfide-bond 11..19
 FT Disulfide-bond 13..31
 FT Disulfide-bond 33..42
 FT Disulfide-bond 50..131
 FT Disulfide-bond 71..113

W Pro-urokinase; thrombolysis; fibrin clot lysis;
W reduced fibrinogenolysis; non-specific plasminogen activation;
W systemic bleeding; mutant Ser175 His187 Gly306.
X
X
X Homo sapiens.

Key Location/Qualifiers
Domain 297..313
Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Disulfide-bond 325..341
Disulfide-bond 352..380

MO9501427-A1.
12-JAN-1995.
28-JUN-1994; 94WO-US07278.
02-JUL-1993; 93US-0087163.
(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
Gurewich V, Liu J;
WPI; 1995-060991/08.
Pro-urokinase mutants - have thrombolytic activity but reduced
fibrinogenolysis activity and non-specific plasminogen activation
Claim 15; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
described in AAR62992-R63008 were derived. These mutants retain the
thrombolytic activity of the wild type protein, useful for the
treatment of thromboembolism, but have a reduced fibrinogenolysis
activity and non-specific plasminogen activation. The mutants can
therefore be used for the lysis of fibrin clots without inducing
systemic bleeding, as can be the case with the wild type protein.
(Updated on 25-MAR-2003 to correct PN field.)
Query Match 100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
121 PLVQECMVHDCADGK 135
121 PLVQECMVHDCADGK 135

25
3
63003 standard; protein; 411 AA.

AC AAR63003;
XX 25-MAR-2003 (updated)
DT 21-SEP-1995 (first entry)
XX
DE Pro-urokinase mutant Ser175 His187 His313.
XX
XX Pro-urokinase; thrombolysis; fibrin clot lysis;
KW reduced fibrinogenolysis; non-specific plasminogen activation;
KW systemic bleeding; mutant Ser175 His187 His313.
OS Homo sapiens.

Key Location/Qualifiers
Domain 297..313
Disulfide-bond 11..19
Disulfide-bond 13..31
Disulfide-bond 33..42
Disulfide-bond 50..131
Disulfide-bond 71..113
Disulfide-bond 102..126
Disulfide-bond 148..279
Disulfide-bond 189..205
Disulfide-bond 197..268
Disulfide-bond 293..362
Disulfide-bond 325..341
Disulfide-bond 352..380
WO9501427-A1.
12-JAN-1995.
28-JUN-1994; 94WO-US07278.
02-JUL-1993; 93US-0087163.
(NEWE-) NEW ENGLAND DEACONESS HOSPITAL.
Gurewich V, Liu J;
WPI; 1995-060991/08.
Pro-urokinase mutants - have thrombolytic activity but reduced
fibrinogenolysis activity and non-specific plasminogen activation
Claim 15; Fig 1; 46pp; English.

AAR62991 is the wild type pro-urokinase, from which the new mutants
described in AAR62992-R63008 were derived. These mutants retain the
thrombolytic activity of the wild type protein, useful for the
treatment of thromboembolism, but have a reduced fibrinogenolysis
activity and non-specific plasminogen activation. The mutants can
therefore be used for the lysis of fibrin clots without inducing
systemic bleeding, as can be the case with the wild type protein.
(Updated on 25-MAR-2003 to correct PN field.)
Query Match 100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
1 SNELHQVPSNCDLNGGTCVSNKYFSNIHWNCNPKKFGGQHCIDKSKTCYEGNGHFYRG 60
61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
61 KASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
121 PLVQECMVHDCADGK 135
121 PLVQECMVHDCADGK 135

25
3
63003 standard; protein; 411 AA.

Db 121 PLVQECMVHDCADGK 135

RESULT 26

AAR63004

ID AAR63004 standard; protein; 411 AA.

XX AC AAR63004;

XX 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX DE Pro-urokinase mutant Ser175 His187 Gly306 Ala313.

XX KW Pro-urokinase; thrombolysis; fibrin clot lysis;

KW reduced fibrinogenolysis; non-specific plasminogen activation;

KW systemic bleeding; mutant Ser175 His187 Gly306 Ala313.

XX OS Homo sapiens.

XX FH Key

FT Domain 297..313 Location/Qualifiers

FT Disulfide-bond 11..19 /note= "flexible loop"

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

PN WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced

FT fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 15; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

CC described in AAR62992-R63008 were derived. These mutants retain the

CC thrombolytic activity of the wild type protein, useful for the

CC treatment of thromboembolism, but have a reduced fibrinogenolysis

CC activity and non-specific plasminogen activation. The mutants can

CC therefore be used for the lysis of fibrin clots without inducing

CC systemic bleeding, as can be the case with the wild type protein.

CC (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 411 AA;

Query Match 100.0%; Score 793; DB 16; Length 411;

Best Local Similarity 100.0%; Pred. No. 1.9e-51;

Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDLNGGTCVSNKVFNSIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60

Db 1 SNELHQVPSNCDLNGGTCVSNKVFNSIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60

Qy 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120

Db 61 KASTDTMGRPCLPWSATVLQOTYHAHRSDALQLGLGKHNCRNPNRRRPPWCYVQVGLK 120

Qy 121 PLVQECMVHDCADGK 135

Db 121 PLVQECMVHDCADGK 135

RESULT 27

AAR63005

ID AAR63005 standard; protein; 411 AA.

XX AC AAR63005;

DT 25-MAR-2003 (updated)

DT 21-SEP-1995 (first entry)

XX DE Pro-urokinase mutant Ser175 His187 Ala300 Ala301 Ala313.

XX KW Pro-urokinase; thrombolysis; fibrin clot lysis;

KW reduced fibrinogenolysis; non-specific plasminogen activation;

KW systemic bleeding; mutant Ser175 His187 Ala300 Ala301 Ala313.

XX OS Homo sapiens.

XX FH Key

FT Domain 297..313 Location/Qualifiers

FT Disulfide-bond 11..19 /note= "flexible loop"

FT Disulfide-bond 13..31

FT Disulfide-bond 33..42

FT Disulfide-bond 50..131

FT Disulfide-bond 71..113

FT Disulfide-bond 102..126

FT Disulfide-bond 148..279

FT Disulfide-bond 189..205

FT Disulfide-bond 197..268

FT Disulfide-bond 293..362

FT Disulfide-bond 325..341

FT Disulfide-bond 352..380

PN WO9501427-A1.

XX 12-JAN-1995.

XX 28-JUN-1994; 94WO-US07278.

XX 02-JUL-1993; 93US-0087163.

XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.

XX Gurewich V, Liu J;

XX WPI; 1995-060991/08.

XX Pro-urokinase mutants - have thrombolytic activity but reduced

FT fibrinogenolysis activity and non-specific plasminogen activation

XX Claim 16; Fig 1; 46pp; English.

XX AAR62991 is the wild type pro-urokinase, from which the new mutants

CC described in AAR62992-R63008 were derived. These mutants retain the

CC thrombolytic activity of the wild type protein, useful for the

CC treatment of thromboembolism, but have a reduced fibrinogenolysis

CC activity and non-specific plasminogen activation. The mutants can

CC therefore be used for the lysis of fibrin clots without inducing

CC systemic bleeding, as can be the case with the wild type protein.

CC (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 411 AA;

```

Query Match          100.0%; Score 793; DB 16; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCETDKSKTCYEGNGHFFYRG 60
Db      1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFTGGQHCETDKSKTCYEGNGHFFYRG 60

QY      61 KASTDTMGRCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNFDRRRRCVQVGLK 120
Db      61 KASTDTMGRCLPWNSATVLQOTYHAHRSDALQLGLGKHNYCRPNFDRRRRCVQVGLK 120

QY      121 PLVQECMVHDCADGK 135
Db      121 PLVQECMVHDCADGK 135

RESULT 28
AAR63006
ID AAR63006 standard; protein; 411 AA.
XX
XC AAR63006:

```

DT	25-MAR-2003	(updated)	
DT	21-SEP-1995	(first entry)	
XX		Pro-urokinase mutant Ser175 His187 His300 Ala301 Ala313.	
DE		XX	
XX		XX	
KW		Pro-urokinase; thrombolysis; fibrin clot lysis;	
KW		reduced fibrinogenolysis; non-specific plasminogen activation;	
KW		systemic bleeding; mutant Ser175 His187 His300 Ala301 Ala313.	
XX		XX	
OS		Homo sapiens.	
XX		XX	
PH		Key	Location/Qualifiers
FT		Domain	297..313
FT			/note= "flexible loop"
FT			11..19
FT		Disulfide-bond	
FT		Disulfide-bond	13..31
FT		Disulfide-bond	33..42

FT	Disulfide-bond	50..131	
FT	Disulfide-bond	71..113	
FT	Disulfide-bond	102..126	
FT	Disulfide-bond	148..279	
FT	Disulfide-bond	189..205	
FT	Disulfide-bond	197..268	
FT	Disulfide-bond	293..362	
FT	Disulfide-bond	325..341	
FT	Disulfide-bond	352..380	
XX	WO9501427-Al.		
PN			
XX			
XX			
PD	12-JAN-1995.		
XX			
XX			
PF	28-JUN-1994;	94WO-US07278.	
XX			
XX			
PR	02-JUL-1993;	93US-0087163.	
XX			
XX	(NEW-) NEW ENGLAND DEACONESS HOSPITAL.		
PA			
XX			
PI	Gurewich V, Liu J;		
XX			
XX			
DR	WPI; 1995-060991/08.		
XX			
XX			
PT	Pro-urokinase mutants - have thrombolytic activity but reduced		
PT	fibrinogenolysis activity and non-specific plasminogen activation		
XX			
XX			
PS	Claim 16; Fig 1; 46pp; English.		
XX			
CC	AA62991 is the wild type pro-urokinase, from which the new mutants		
CC	described in AA62992-R63008 were derived. These mutants retain the		
CC	thrombolytic activity of the wild type protein, useful for the		
CC	treatment of thromboembolism, but have a reduced fibrinogenolysis		

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 6.98276 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-4
Perfect score: 793
Sequence: 1 SNELHQVPSNCDCLNGGTGV.....QVGLKPLVQECMVHDCADGK 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.76.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	793	100.0	431	1 UKHU	u-plasminogen acti
2	713.5	90.0	433	1 UKBAY	u-plasminogen acti
3	608	76.7	442	1 UKPG	u-plasminogen acti
4	579	73.0	433	1 JN0560	u-plasminogen acti
5	571	72.0	432	1 S18932	u-plasminogen acti
6	548	69.1	433	1 UPLMS	u-plasminogen acti
7	335.5	42.3	477	1 A34369	t-plasminogen acti
8	335.5	42.3	477	2 JS0598	t-plasminogen acti
9	334.5	42.2	434	1 A35005	u-plasminogen acti
10	328.5	41.4	291	2 J38098	t-plasminogen acti
11	328.5	41.4	431	2 JS0599	t-plasminogen acti
12	328.5	41.4	562	1 UKHUT	t-plasminogen acti
13	321.5	40.5	559	1 A25029	t-plasminogen acti
14	315.5	39.8	559	1 A29941	t-plasminogen acti
15	310.5	39.2	477	2 JS0597	t-plasminogen acti
16	264.5	33.4	558	2 JC5878	plasma hyaluronan-
17	260.5	32.8	655	1 A46688	hepatocyte growth
18	257.5	32.5	560	1 JC4795	plasma hyaluronan-
19	257	32.4	603	2 S28941	coagulation factor
20	233	29.4	615	1 KFHU12	coagulation factor
21	223	28.1	394	2 JS0600	t-plasminogen acti
22	217	27.4	593	2 S45281	coagulation factor
23	169	21.3	685	1 A48289	neurotrophic recep
24	162.5	20.5	4548	1 S00657	apoptosis(a) (SC
25	161	20.3	1420	2 A32869	apolipoprotein(a)
26	160	20.2	123	2 C61545	plasmin (EC 3.4.21
27	159	20.1	120	2 E61545	plasmin (EC 3.4.21
28	158.5	20.0	937	2 A45082	neurotrophic recep
29	157.5	19.9	460	2 B61545	plasmin (EC 3.4.21

30	156	19.7	812	1 PLBO	plasmin (EC 3.4.21
31	155	19.5	790	1 PLPG	plasmin (EC 3.4.21
32	154	19.4	89	2 A60140	plasmin (EC 3.4.21
33	154	19.4	810	1 PLHU	plasmin (EC 3.4.21
34	153.5	19.4	810	2 I46260	plasmin (EC 3.4.21
35	152.5	19.2	810	2 B30848	plasmin (EC 3.4.21
36	152	19.2	169	2 A40522	plasmin (EC 3.4.21
37	148	18.7	943	2 B45082	neurotrophic recep
38	147	18.5	812	1 PLMS	plasmin (EC 3.4.21
39	144.5	18.2	806	2 T18840	hypothetical prote
40	144	18.2	711	1 A47136	macrophage-stimula
41	141.5	17.8	455	2 A61545	plasmin (EC 3.4.21
42	141.5	17.8	728	1 JH0579	hepatocyte growth
43	141	17.8	411	2 I51285	hepatocyte growth
44	141	17.8	710	1 I51283	hepatocyte growth
45	139	17.5	716	1 A40332	macrophage-stimula

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human

N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminog

N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a

in form

C;Species: Homo sapiens (man)

C;Date: 17-Dec-1992 #sequence revision 04-Dec-1986 #text change 15-Sep-2000

C;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A

R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.

Nucleic Acids Res. 13, 2759-2771, 1985

A;Title: The human urokinase-plasminogen activator gene and its promoter.

A;Reference number: A00931; MUID:85215647; PMID:2987867

A;Accession: A00931

A;Molecule type: DNA

A;Residues: 1-431 <RIC>

A;Cross-references: GB:X02419; NID:g37601; PIDN:CAA26268.1; PID:g1834524

A;Note: the authors translated the codon ATG for residue 214 as Ile

R;Nagamine, Y.; Pearson, D.; Grattan, M.

Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin

A;Reference number: I52209; MUID:86050639; PMID:3933505

A;Accession: I52209

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 145-161 <NAG1>

A;Cross-references: GB:K03027; NID:g340174; PIDN:AAA61257.1; PID:g340175

R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama

Gene 36, 183-188, 1985

A;Title: Molecular cloning of cDNA coding for human preprourokinase.

A;Reference number: J70102; MUID:86056954; PMID:2415429

A;Accession: J70102

A;Molecule type: mRNA

A;Residues: 1-213, 'I', 215-431 <NAG2>

A;Cross-references: GB:K03226; NID:g340155; PIDN:AA097138.1; PID:g340158; GB:D00244; N

R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) +

A;Reference number: A37561; MUID:84272706; PMID:6589620

A;Accession: A37561

A;Molecule type: mRNA

A;Residues: 66-431 <VER>

A;Cross-references: GB:D00244; NID:g220138

R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Else.

DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr

A;Reference number: I38102; MUID:85203359; PMID:3888571

A;Accession: I38102

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>

A;Cross-references: EMBL:X02760; NID:g35297; PIDN:CAA26535.1; PID:g35298

R;Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, Biochim. Biophys. Acta 1293, 83-89, 1996
A;Title: Characterization of single chain urokinase-type plasminogen activator with a novel A;Reference number: S65783; MUID:96186279; PMID:8652631
A;Accession: S65783
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
A;Cross-references: EMBL:D11143; NID:G1311467; PIDN:BAA01919.1; PID:G1199928
R;Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frank, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
A;Title: The primary structure of high molecular mass urokinase from human urine. A;Reference number: A37562; MUID:83055084; PMID:6754569
A;Accession: A37562
A;Molecule type: protein
A;Residues: 21-177 <GUN>
R;Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O. Eur. J. Biochem. 125, 251-257, 1982
A;Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary A;Reference number: A37563; MUID:83003608; PMID:6749491
A;Accession: A37563
A;Molecule type: protein
A;Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>
R;Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frank, E.; Flohe, L. Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
A;Title: The complete amino acid sequence of low molecular mass urokinase from human urine A;Reference number: A37564; MUID:83055099; PMID:6754572
A;Accession: A37564
A;Molecule type: protein
A;Residues: 159-410 <STE>
R;Kentzer, E.J.; Boko, A.; Menon, G.; Sarin, V.K. Biochem. Biophys. Res. Commun. 171, 401-406, 1990
A;Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant A;Reference number: A35689; MUID:90365737; PMID:2393398
A;Accession: A35689
A;Molecule type: protein
A;Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
A;Note: identification of a fucose and attempt to determine its attachment site
R;Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzman Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
A;Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line A;Reference number: A36697; MUID:91097529; PMID:2125213
A;Accession: A36697
A;Molecule type: protein
A;Residues: 21-34 <RAB>
R;Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M. Submitted to the Brookhaven Protein Data Bank, July 1993
A;Reference number: A51255; PDB:1KDU
A;Contents: annotation: conformation and disulfide bond assignments by (1)H-NMR, residue R;Li, X.; Smith, R.A.G.; Dobson, C.M. Biochemistry 31, 9562-9571, 1992
A;Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of A;Reference number: A44375; MUID:93003110; PMID:1327118
A;Contents: annotation: conformation and disulfide bond assignments by (1)H-NMR
R;Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak, submitted to the Brookhaven Protein Data Bank, January 1994
A;Reference number: A66822; PDB:1URK
A;Contents: annotation: conformation and disulfide bond assignments by (1)H-NMR, residue R;Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.; submitted to the Brookhaven Protein Data Bank, July 1995
A;Reference number: A66058; PDB:1LW
A;Contents: annotation: X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
C;Comment: this enzyme is found in urine in a high molecular mass form, consisting of A C;Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a C;Genetics:
A;Gene: GDB:PLAU
A;Cross-references: GDB:119497; OMIM:191840
A;Map position: 10Q24-10Q24
A;Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
C;Function:
A;Description: proteolytically activates plasminogen
A;Pathway: fibrinolysis
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C;Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted <M
F;21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M
F;31-62/Domain: EGF homology <EGF>
F;70-151/Domain: kringle homology <KRG>
F;156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental <F
F;179-431/Product: urokinase-type plasminogen activator chain B #status experimental <F
F;31-33, 33-51, 51-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-
F;38/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;178-179/Cleavage site: Lys-Ile (plasmin) #status experimental
F;224, 275, 376/Active site: His, Asp, Ser #status experimental
F;322/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 100.0%; Score 793; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.2e-64;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFYRG 80
Qy 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 140
Qy 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155
RESULT 2
UKBAY
u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon
C;Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C;Accession: S14687; S08651
R;Au, Y.F.T.; Wang, T.W.; Clowes, A.W. Nucleic Acids Res. 18, 3411, 1990
A;Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasmino
A;Reference number: S14687; MUID:90287734; PMID:2113276
A;Accession: S14687
A;Molecule type: mRNA
A;Residues: 1-433 <AU>
A;Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131
C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-176/Product: plasminogen activator chain A #status predicted <ACH>
F;30-61/Domain: EGF homology <EGF>
F;169-150/Domain: kringle homology <KRG>
F;178-433/Product: plasminogen activator chain B #status predicted <BCH>
F;178-421/Domain: trypsin homology <TRY>
F;167-298, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F;223, 274, 378/Active site: His, Asp, Ser #status predicted
F;324/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 90.0%; Score 713.5; DB 1; Length 433;
Best Local Similarity 91.9%; Pred. No. 4.8e-57;
Matches 124; Conservative 4; Mismatches 6; Indels 1; Gaps 1;
Qy 1 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFYRG 60
Db 21 SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCETDKSKTCYEGNGHFYRG 79
Qy 61 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 120
Db 81 KASDTMTGRPCLPWNSATVLQQTTHAHRSDALQLGLGKHNYCRNPNRRPWCYVQVGLK 139
Qy 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 154

RESULT 12

UKHTT

N;Alternate names: t-PA; tissue plasminogen activator
C;Species: Homo sapiens (man)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1983 #text_change 08-Dec-2000
C;Accession: A94004; A23529; J070562; A93293; S02125; A91343; A93951; A91322; A54645; I60
R;NY, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A;Title: The structure of the human tissue-type plasminogen activator gene: correlation
A;Reference number: A94004; MUID:84298137; PMID:6089198
A;Accession: A94004
A;Molecule type: DNA
A;Residues: 1-562 <NYT>
A;Cross-references: GB:I00141
A;Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
P;Friezen Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A;Title: The human tissue plasminogen activator gene.
A;Reference number: A23529; MUID:86196143; PMID:3009482
A;Accession: A23529
A;Molecule type: DNA
A;Residues: 1-562 <DBG>
A;Cross-references: GB:K03021; NID:G339817; PIDN:AAA98809.1; PID:G339818
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A;Title: Purification and characterization of tissue plasminogen activator secreted by H
A;Reference number: J070562; MUID:91291340; PMID:1368681
A;Accession: J070562
A;Molecule type: mRNA
A;Residues: 31-562 <ITA>
A;Cross-references: DBJ:D01096; NID:9220128; PIDN:BAA00881.1; PID:G441174
A;Experimental source: embryonic lung fibroblast IMR-90 cells
A;Note: part of this sequence, including the amino end of the mature protein, was confir
P;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vohar, G.A.; Ward, C.A.; Bennett
Nature 301, 214-221, 1983
A;Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Esche
A;Reference number: A93293; MUID:83115262; PMID:6337343
A;Accession: A93293
A;Molecule type: mRNA
A;Residues: 1-562 <PEN>
A;Cross-references: GB:I00141
A;Experimental source: melanoma cells
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A;Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human fe
A;Reference number: S02125; MUID:86262579; PMID:3133640
A;Accession: S02125
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-562 <SAS>
A;Cross-references: EMBL:X07393; NID:G37243; PIDN:CAA30302.1; PID:G37244
A;Experimental source: fetal lung cells
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; Ma
FEBS Lett. 189, 145-149, 1985
A;Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A;Reference number: A91343; MUID:85285620; PMID:3896853
A;Accession: A91343
A;Molecule type: mRNA
A;Residues: 1-38,'G',86-433,'E',435-562 <KAG>
A;Experimental source: Detroit 562 cells; ATCC 138
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A;Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen activ
A;Reference number: A93951; MUID:83169656; PMID:6572897
A;Accession: A93951
A;Molecule type: mRNA
A;Residues: 251-358 <EDL>
A;Experimental source: melanoma cells
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A;Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived am
differences.

A;Reference number: A90488; MUID:85000468; PMID:6433976
A;Contents: annotation; melanoma cells, partial sequence of residues 36-562, active an
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984
A;Title: Differences between uterine and melanoma forms of tissue plasminogen activator;
A;Reference number: A91322; MUID:84158956; PMID:6538514
A;Accession: A91322
A;Molecule type: protein
A;Residues: 33-45;311-320 <POH>
A;Experimental source: uterus
A;Note: in the uterus, cleavage of the activation peptide may also occur after 38-Gln
R;van Zonneveld, A.J.; Veerman, H.; Pannekoek, H.
J. Biol. Chem. 261, 14214-14218, 1986
A;Reference number: A37567; MUID:97033611; PMID:3021732
A;Contents: annotation; fibrin binding site
R;Verheijen, J.H.; Caspers, M.P.M.; Chang, G.T.G.; de Munk, G.A.W.; Pouwels, P.H.; Eng
EMBO J. 5, 3525-3530, 1986
A;Title: Involvement of finger domain and kringle 2 domain of tissue-type plasminogen
A;Reference number: A37568; MUID:87161761; PMID:3030730
A;Contents: annotation; fibrin binding site
R;Dodd, I.; Nunn, B.; Robinson, J.H.
Thromb. Haemost. 59, 523-528, 1988
A;Title: Isolation, identification and pharmacokinetic properties of human tissue-type
A;Reference number: A60902; MUID:89044681; PMID:3142086
R;Harris, T.J.R.; Patel, T.; Marston, F.A.O.; Little, S.; Emtage, J.S.; Opdenakker, G.
Mol. Biol. Med. 3, 279-292, 1986
A;Title: Cloning of cDNA coding for human tissue-type plasminogen activator and its ex
A;Reference number: A54645; MUID:86284200; PMID:3090401
A;Accession: A54645
A;Molecule type: mRNA
A;Residues: 1-562 <HAR>
A;Cross-references: GB:M15518; NID:G190031; PIDN:AAA60111.1; PID:G190032
A;Note: parts of this sequence were confirmed by peptide sequencing
R;Reddy, V.B.; Garramone, A.J.; Sasak, H.; Wei, C.
DNA 6, 461-472, 1987
A;Title: Expression of human uterine tissue-type plasminogen activator in mouse cells
A;Reference number: 160110; MUID:88054470; PMID:2824147
A;Accession: I60110
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-562 <RES>
A;Cross-references: GB:M18182; NID:G340176; PIDN:AAA36800.1; PID:G340177
R;Fisher, R.; Waller, E.K.; Grossi, G.; Thompson, D.; Tizard, R.; Schleuning, W.D.
J. Biol. Chem. 260, 11223-11230, 1985
A;Title: Isolation and characterization of the human tissue-type plasminogen activator
A;Reference number: 155232; MUID:85289338; PMID:3161893
A;Accession: 155232
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-36 <RE2>
A;Cross-references: GB:M11890; NID:G339837; PIDN:AAA61213.1; PID:G339839
C;Comment: Cleavage by plasmin or trypsin produces two chains held together by a singl
C;Comment: t-PA converts plasminogen to plasmin by hydrolyzing a single Arg-Val bond.
C;Comment: t-PA binds chain A of fibrin by kringle 2 and the fibronectin type I repeat
C;Genetics:
A;Gene: GDB:PLAT
A;Cross-references: GDB:119496; OMIM:173370
A;Map position: 8p12-8p12
A;Introns: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2; 297/1; 362/2; 408/1; 455/3; 5
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-32/Domain: propeptide #status predicted <PRO>
F;33-562/Product: t-plasminogen activator #status experimental <MAT>
F;33-310/Product: t-plasminogen activator chain A #status experimental <ACH>
F;41-78/Domain: fibronectin type I repeat homology <IFI>
F;86-119/Domain: EGF homology <EGF>
F;127-208/Domain: kringle homology <KR1>
F;215-296/Domain: kringle homology <KR2>
F;311-562/Product: t-plasminogen activator chain B #status experimental <BCH>
F;311-556/Domain: trypsin homology <TRY>
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203,215-296,236-278,267-291,299

F;152,483/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;210-311/Binding site: carbohydrate (Asn) (covalent) (partial) #status experimental
F;310-311/Cleavage site: Arg-Ile (plasmin, trypsin) #status experimental
F;357,406/Active site: His, Asp #status predicted
F;513/Active site: Ser #status experimental

Query Match 41.4%; Score 328.5; DB 1; Length 562;
Best Local Similarity 46.3%; Pred. No. 3.3e-22;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

Qy 3 ELHQP-SNCD---CLNGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGTCQQALYFSDF-VCQPEGAGKCEIDTTRATCYEDGIGSY 135
Qy 59 RGKASTDTMGRCPLPWNATVLQOTYHAHRSDALQLGLGHNHNYCRNPNRRPWCYVQVG 118
Db 136 RGTWSTAESGAECINWSSALAQKPYSGRRPDALRLGLGHNHNYCRNPNRRDLSKPCYVFKA 195
Qy 119 LKPLVQECMVHDCADG 134
Db 196 GKYTSSEFCSTPACSEG 211

RESULT 13
A35029
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A35029; A31597
R;Feng, P.; Ohlsson, M.; Ny, T.
J. Biol. Chem. 285, 2022-2027, 1990
A;Title: The structure of the TARA-less rat tissue-type plasminogen activator gene. Sped
A;Reference number: A35029; MUID:90130448; PMID:2105315
A;Accession: A35029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-559 <FEN>
A;Cross-references: GB:M31197; NID:G207429; PID:AAA42261.1; PID:G207431; GB:J05226
R;Ny, T.; Leonardsson, G.; Hsueh, A.J.W.
DNA 7, 671-677, 1988
A;Title: Cloning and characterization of a cDNA for rat tissue-type plasminogen activator
A;Reference number: A31597; MUID:99170114; PMID:3148445
A;Accession: A31597
A;Molecule type: mRNA
A;Residues: 1-379, 'K', 381-559 <NYT>
A;Cross-references: GB:M23697; NID:G530159; PID:AAA41812.1; PID:G530160
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297-4
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 40.5%; Score 321.5; DB 1; Length 559;
Best Local Similarity 45.6%; Pred. No. 1.4e-21;
Matches 62; Conservative 14; Mismatches 55; Indels 5; Gaps 2;

Qy 3 ELHQP-----SNCDLNGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFY 58
Db 74 QCHSVPVKSCSEPRCFNGTCQQALYFSDF-VCQCPDGFVGRKCDIDTRATCFEGQGIT 132
Qy 59 RGKASTDTMGRCPLPWNATVLQOTYHAHRSDALQLGLGHNHNYCRNPNRRPWCYVQVG 118

Db 133 RGTWSTAENGAEICINWSSALSQKPYSAARRPNAIKLGLGHNHNYCRNPNRRDVKPCYVFKA 192
Qy 119 LKPLVQECMVHDCADG 134
Db 193 GKYTTEFCSTPACPKG 208

RESULT 14
A29941
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29941; S48205; S48207; S48206
R;Richles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator r
A;Reference number: A29941; MUID:88087303; PMID:2826484
A;Accession: A29941
A;Molecule type: mRNA
A;Residues: 1-559 <RIC>
A;Cross-references: GB:J03520; NID:G202109; PID:AAA40470.1; PID:G202110
R;Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076; PMID:7523120
A;Accession: S48205
A;Molecule type: protein
A;Residues: 33-37, 'X', 39-40 <LIJ>
A;Accession: S48207
A;Molecule type: protein
A;Residues: 309-316 <LI2>
A;Accession: S48206
A;Molecule type: protein
A;Residues: 33-37, 'X', 39-40 <LIW>
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-205,145-187,176-200,213-294,234-276,265-289,297
F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 39.8%; Score 315.5; DB 1; Length 559;
Best Local Similarity 44.5%; Pred. No. 4.9e-21;
Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;

Qy 3 ELHQP-----SNCDLNGTCVSNKYFSNIHWCNCPKFGQHCEIDKSKTCYEGNGHFY 58
Db 74 QCHSVPVKSCSEPRCFNGTCQQALYFSDF-VCQCPDGFVGRKCDIDTRATCFEGQGIT 132
Qy 59 RGKASTDTMGRCPLPWNATVLQOTYHAHRSDALQLGLGHNHNYCRNPNRRPWCYVQVG 118
Db 133 RGTWSTAENGAEICINWSSVLKPYNARRPNAIKLGLGHNHNYCRNPNRRDLKPCYVFKA 192
Qy 119 LKPLVQECMVHDCADG 135
Db 193 GKYTTEFCSTPACPKG 209

RESULT 15
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmondus rotundus (common vampire bat)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 R;Xraetzschar, J.; Haendler, B.; Langer, G.; Boidel, W.; Bringmann, P.; Alagon, A.; Dor
 Gene 105, 229-237, 1991
 A>Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A;Reference number: JS0597; MUID:92039036; PMID:1937019
 A;Accession: JS0597
 A;Molecule type: mRNA
 A;Residues: 1-477 <KRA>
 A;Cross-references: GB:M63987; NID:g166070; PIDN:AAA31591.1; PID:g166071
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <1FA>
 F;87-120/Domain: EGF homology <EGF>
 F;128-209/Domain: kringle homology <KRG>
 F;226-471/Domain: trypsin homology <TRY>
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-4
 F;153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
 F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 39.28; Score 310.5; DB 2; Length 477;
 Best Local Similarity 45.58; Pred. No. 1.2e-20;
 Matches 60; Conservative 14; Mismatches 53; Indels 5; Gaps 3;

Qy 5 HQVPSN-CD---CLNGGTCVSNKYFSNIHNCNPKFGQHCIDKSKTCYEGNGHYRFG 60
 Db 80 HTVPVNSCEPRFCNGGTCQAVFSDP-VQCQAGTYGKCEVDTRATCYEGGQVIYRG 138

Qy 61 KASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLK 120
 Db 139 TWSTAERVEICINWSSILLTRTYNGRMPDAFNLGLGNHNYCRNPNGAPKFCWYIRAGK 198

Qy 121 PLVQECMVHDC 132

Db 199 FTSESCVPVCS 210

RESULT 16
 Plasma hyaluronan-binding protein precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
 C;Accession: JC5878
 R;Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, N.
 Biol. Pharm. Bull. 20, 1127-1130, 1997
 A>Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-bin
 A;Reference number: JC5878; MUID:98065239; PMID:9401717
 A;Accession: JC5878
 A;Molecule type: mRNA
 A;Residues: 1-558 <HAS>
 C;Comment: This protein acts as serine protease.
 C;Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsi
 F;1-23/Domain: signal sequence #status predicted <SIG>
 F;24-311/Product: plasma hyaluronan-binding protein large chain #status predicted <MATL>
 F;75-106/Domain: EGF homology <EG1>
 F;113-145/Domain: EGF homology <EG2>
 F;152-185/Domain: EGF homology <EG3>
 F;192-274/Domain: kringle homology <KRI>
 F;312-558/Product: plasma hyaluronan-binding protein small chain #status predicted <MATS
 F;312-548/Domain: trypsin homology <TRY>

Query Match 33.44; Score 264.5; DB 2; Length 558;
 Best Local Similarity 42.58; Pred. No. 2e-16;
 Matches 51; Conservative 18; Mismatches 48; Indels 3; Gaps 3;

Qy 13 CLNGGTCVSNKYFSNIHNCNPKFGQHCIDKSKTCYEGNGHYRFGKASTDTMGRCPL 72
 Db 157 CONGGVCSRRHRRSRF-TCACPDQYKGFCEIGPD-DCYVGDGYSYRGKSVKTVNQNPCL 214

Qy 73 PWSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCYVQVGLKPLVQECMVHDC 131
 Db 215 YNNSHLLQETYNMFMEDEATHGIAEHNFPCRDGDHPCWCFVKVNSEKVKWEYCDVTVC 274

RESULT 17
 A46688
 hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
 C;Species: Homo sapiens (man)
 C;Date: 21-Sep-1993 #sequence_revision 25-Aug-1995 #text_change 08-Dec-2000
 C;Accession: A46688
 R;Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
 J. Biol. Chem. 268, 10024-10028, 1993
 A>Title: Molecular cloning and sequence analysis of the cDNA for a human serine protea
 d coagulation factor XII.
 A;Reference number: A46688; MUID:93252878; PMID:7683665
 A;Accession: A46688
 A;Molecule type: mRNA
 A;Residues: 1-655 <MTY>
 A;Cross-references: DDBJ:D14012; NID:g219680; PIDN:BAR03113.1; PID:g219681
 A;Experimental source: liver (mRNA); serum (protein)
 A;Note: sequence extracted from NCBI backbone (NCBIN:131227, NCBIPI:131228)
 C;Note: parts of the sequence, including the amino ends of the heavy and light chains,
 C;Genetics:
 A;Gene: GDB:HGFAC; HGFA; HGFAP
 A;Cross-references: GDB:9954514
 A;Map position: 4p16-4p16
 C;Function:
 A;Description: activates hepatocyte growth factor by specific proteolytic cleavage
 A;Pathway: tissue repair and regeneration
 C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homolog
 C;Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
 F;1-34/Domain: signal sequence #status predicted <SIG>
 F;108-148/Domain: fibronectin type II repeat homology <1F2>
 F;164-197/Domain: EGF homology <EG1>
 F;202-237/Domain: fibronectin type I repeat homology <1F1>
 F;245-278/Domain: EGF homology <EG2>
 F;286-367/Domain: kringle homology <KRG>
 F;373-407/Product: hepatocyte growth factor activator light chain #status experimental
 F;408-655/Product: hepatocyte growth factor activator heavy chain #status experimental
 F;408-641/Domain: trypsin homology <TRY>
 F;40,48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-
 F;447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 32.88; Score 260.5; DB 1; Length 655;
 Best Local Similarity 43.78; Pred. No. 5.1e-16;
 Matches 52; Conservative 8; Mismatches 42; Indels 17; Gaps 2;

Qy 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPKFGQHCIDKSKTCYEGNGH 56
 Db 242 HTACLSFCLNGGTC-----HLIVATGTTVCAPPGPAGRLCNIEPDERFLNGT 292

Qy 57 FYRGKASTDTMGRCPLPNSATVLOQTYHAHRSALQGLGKHNYCRNPDRRRPWCY 115
 Db 293 GYRGVASTASAGLSCLAWSNLLYQELHVDVSGAALLGLGPHAYCRNPDRPWCY 351

RESULT 18
 JC4795
 plasma hyaluronan-binding protein precursor - human
 N;Alternate names: hepatocyte growth factor activator-like protein; PHBP
 N;Contains: serine proteinase (EC 3.4.21.-)
 C;Species: Homo sapiens (man)
 C;Date: 15-Oct-1995 #sequence_revision 16-Aug-1996 #text_change 19-Jul-2002
 C;Accession: JC4795
 R;Choi-Miura, N.H.; Tobe, T.; Sumiya, J.; Nakano, Y.; Sano, Y.; Mazda, T.; Tomita, M.
 J. Biochem. 119, 1157-1165, 1996
 A>Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP
 r activator.
 A;Reference number: JC4795; MUID:96425001; PMID:8827452
 A;Accession: JC4795
 A;Molecule type: mRNA

Thu Dec 4 06:30:58 2003

```
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-372,373-615/Product: coagulation factor XIIa, alpha form #status experimental <A12>
F;47-88/Domain: fibronectin type II repeat homology <FB2>
F;98-130/Domain: EGF homology <EG1>
F;135-170/Domain: fibronectin type I repeat homology <1F1>
F;178-209/Domain: EGF homology <EG2>
F;217-295/Domain: kringle homology <KR>
F;298-356/Region: proline-rich
F;354-362,373-615/Product: coagulation factor XIIa, beta form #status experimental <B12>
F;373-609/Domain: trypsin homology <TRY>
F;98-110,104-119,121-130,135-163,161-170,178-189,183-198,200-209,217-295,238-277,266-290
F;109/Binding site: carbohydrate (Thr) (covalent) #status experimental
F;249,433/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;239,305,328,329,337/Binding site: carbohydrate (Thr) (covalent) #status predicted
F;308/Binding site: carbohydrate (Ser) (covalent) #status predicted
F;412,461,563/Active site: His, Asp, Ser #status predicted

Query Match          29.4%; Score 233; DB 1; Length 615;
Best Local Similarity 42.3%; Pred. No. 1.5e-13;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

Qy 13 CLNGGTCVSNKYFSNIHWCNPKFGQHCIDKSKTCYEGNGHFGYRGKASTDTMGRLPCL 72
Db 183 CLHGGRCL-----VEGHLRCHCPVGVTGPFCDVDTKASCDYGRGLSYRGLARTTLGAPCQ 239

Qy 73 PWSATVLQQTYY-HAHRSDALQGLGKHNKNCNPDNRRPWCYV 115
Db 240 PWAS-----EATYVNTAEQVNWGLGHAFCRNPDNRRPWCYV 279

RESULT 21
JS0600
t-plasminogen activator (EC 3.4.21.68) gamma precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0600
R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0600
A;Molecule type: mRNA
A;Residues: 1-394 <KRA>
A;Cross-references: GB:M63990; NID:g166078; PIDN:AAA311595.1; PID:g166079
A;Note: the authors translated the codon ATC for residue 75 as Thr
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-394/Product: plasminogen activator gamma #status predicted <PLA>
F;45-126/Domain: kringle homology <KR>
F;143-388/Domain: trypsin homology <TRY>
F;45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F;142-143/Cleavage site: His-Ser (plasma) #status predicted
F;189,238,345/Active site: His, Asp, Ser #status predicted
F;135/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          28.1%; Score 223; DB 2; Length 394;
Best Local Similarity 44.3%; Pred. No. 8.1e-13;
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0;

Qy 45 DSKSKTCYEGNGHFGYRGKASTDTMGRLPWSATVLQQTYYHAHRSDALQGLGKHNKNCN 104
Db 40 DPHATYKDDQGVYTGVTSTSSGQAQCNWNNLLIRTYGRMPPEAVKVLGNGHNKNCN 99

Qy 105 PDNRRRPWCYVQVGLKPLVQECWHDCA 132
Db 100 PDGASKPWCYVVKARFTSESQVPCS 127

RESULT 22
S45281
```

```
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N;Alternate names: Hageman factor (activated)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C;Accession: S45281; A61329
R;Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A;Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): coi
A;Reference number: S45281; MUID:94242782; PMID:8186251
A;Accession: S45281
A;Molecule type: mRNA
A;Residues: 1-593 <SHI>
A;Cross-references: GB:S70164
A;Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as
is, and ATC for residue 505 as Leu
R;Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A;Title: Isolation and characterization of bovine factor XII (Hageman factor).
A;Reference number: A61329; MUID:77182112; PMID:861210
A;Accession: A61329
A;Molecule type: protein
A;Residues: 10-16,'X',18-19;525-550 <FUJ>
C;Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homol
C;Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma;
F;37-78/Domain: fibronectin type II repeat homology <1F2>
F;86-120/Domain: EGF homology <EGF>
F;125-160/Domain: fibronectin type I repeat homology <FB1>
F;207-287/Domain: kringle homology <KR>
F;350-587/Domain: trypsin homology <TRY>
F;541/Active site: Ser #status predicted

Query Match          27.4%; Score 217; DB 2; Length 593;
Best Local Similarity 35.7%; Pred. No. 4e-12;
Matches 46; Conservative 15; Mismatches 58; Indels 10; Gaps 4;

Qy 6 QVPSNCCLNGGTCVSNKYFSNIHWCNPKFGQHCIDKSKTCYE--GNGHFYRGKAS 63
Db 166 QVCTNPLNGDSCLQAE---GHLRCPSPFAGRLCDVDLKASCYDDRDGLSYRGWAG 222

Qy 64 TDTMGRLPWSATVLQQTYY-HAHRSDALQGLGKHNKNCNPDNRRPWCYVQVGLKPL 122
Db 223 TTLGAPCQSWAS----EATYVNTAEQVNWGLGHAFCRNPDNTRPWCPIWKGDRLS 278

Qy 123 VQECWHDCA 131
Db 279 WNYCRLAPC 287

RESULT 23
A48289
neurotrophic receptor ror precursor - fruit fly (Drosophila melanogaster)
N;Alternate names: trk-related receptor
N;Contains: protein-tyrosine kinase (EC 2.7.1.112)
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Nov-2000
C;Accession: A48289
R;Wilson, C.; Goberdhan, D.C.I.; Steller, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 7109-7113, 1993
A;Title: Drosophila neurotrophic receptor gene, encodes a Drosophila homolog of
A;Reference number: A48289; MUID:93348222; PMID:8394009
A;Accession: A48289
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-685 <WIL>
A;Cross-references: GB:L20297; NID:g348103; PIDN:AAA28860.1; PID:g348104
C;Genetics:
A;Gene: FlyBase:bsk
A;Cross-references: FlyBase:FBgn0010407
C;Superfamily: Drosophila neurotrophic receptor ror; kringle homology; protein kinase
C;Keywords: ATP; glycoprotein; kringle; phosphotransferase; transmembrane protein; tyr
F;237-310/Domain: kringle homology <KR>
F;314-338/Domain: transmembrane #status predicted <TM1>
```


A>Status: preliminary; translated from GB/EMBL/DBDB

A:Molecule type: DNA

A:Residues: 1-16 <R3>

A:Cross-references: GB:M86877; NID:gl78780; PIDN:AA849909.1; PID:9553185

A>Note: apo(a) gene 1 (nomenclature of reference I52415)

A:Accession: I65286

A>Status: preliminary; translated from GB/EMBL/DBDB

A:Molecule type: DNA

A:Residues: 1-16 <R4>

A:Cross-references: GB:M86878; NID:gl78782; PIDN:AA851749.1; PID:9553186

C:Genetics:

A:Gene: GDB:LPA

A:Cross-references: GDB:120699; OMIM:152200

A:Map position: 6q26-6q27

A>Note: several genes closely linked on chromosome 6 are identical in the first coding rs of kringle repeats

C:Superfamily: apolipoprotein(a); kringle homology; trypsin homology

C:Keywords: hydrolase; kringle; lipid binding; lipoprotein; serine proteinase

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-4548/Product: apolipoprotein(a) #status experimental <MAT>

F:28-105/Domain: kringle homology <KRI>

F:142-219/Domain: kringle homology <KR2>

F:256-333/Domain: kringle homology <KR3>

F:370-447/Domain: kringle homology <KR4>

F:484-561/Domain: kringle homology <KR5>

F:598-675/Domain: kringle homology <KR6>

F:712-789/Domain: kringle homology <KR7>

F:826-903/Domain: kringle homology <KR8>

F:940-1017/Domain: kringle homology <KR9>

F:1054-1131/Domain: kringle homology <KR10>

F:1168-1245/Domain: kringle homology <KR11>

F:1282-1359/Domain: kringle homology <KR12>

F:1396-1473/Domain: kringle homology <KR13>

F:1510-1587/Domain: kringle homology <KR14>

F:1624-1701/Domain: kringle homology <KR15>

F:1738-1815/Domain: kringle homology <KR16>

F:1852-1929/Domain: kringle homology <KR17>

F:1966-2043/Domain: kringle homology <KR18>

F:2080-2157/Domain: kringle homology <KR19>

F:2194-2271/Domain: kringle homology <KR20>

F:2308-2385/Domain: kringle homology <KR21>

F:2422-2499/Domain: kringle homology <KR22>

F:2536-2613/Domain: kringle homology <KR23>

F:2650-2727/Domain: kringle homology <KR24>

F:2764-2841/Domain: kringle homology <KR25>

F:2878-2955/Domain: kringle homology <KR26>

F:2992-3069/Domain: kringle homology <KR27>

F:3106-3183/Domain: kringle homology <KR28>

F:3220-3297/Domain: kringle homology <KR29>

F:3334-3411/Domain: kringle homology <KR30>

F:3448-3525/Domain: kringle homology <KR31>

F:3562-3639/Domain: kringle homology <KR32>

F:3676-3753/Domain: kringle homology <KR33>

F:3782-3859/Domain: kringle homology <KR34>

F:3896-3973/Domain: kringle homology <KR35>

F:4010-4087/Domain: kringle homology <KR36>

F:4124-4201/Domain: kringle homology <KR37>

F:4228-4307/Domain: kringle homology <KR38>

F:4328-4541/Domain: trypsin homology <TRY>

Query Match 20.5%; Score 162.5; DB 1; Length 4548;
Best Local Similarity 33.1%; Pred. No. 1.8e-06;
Matches 41; Conservative 12; Mismatches 46; Indels 25; Gaps 7;

Qy 27 NIHW--CN---CP-----KFGGQHCHIDKSKT---CYEGNGHFYRGKASTDTMGR 69
Db 3742 NVRWECNLTQCPVTESSVLATSAVSQAPTEQSPQTQDCYHGDGQS YRGSPFSITVTGR 3801
Qy 70 PCLPWNASATVLQOYTHAHRSDALQLGLKHNVCNPNDRRRPWCYVQGLKPLV--QECM 127
Db 3802 TCQSWSSMT---PHWHORTTEYPNGGLTRNVCNPNDAEIRPWCYT--MDPSVRWEYCN 3855
Qy 128 VHDC 131

A;Note: it is uncertain whether Met-1 or Met-8 is the initiator
R;Schaller, J.; Moser, P.W.; Dannegger-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Ri-
Eur. J. Biochem. 149, 267-278, 1985
A;Title: Complete amino acid sequence of bovine plasminogen. Comparison with human pla-
A;Reference number: A25835; MUID: 85203906; PMID: 3846532
A;Accession: A25835
A;Molecule type: protein
A;Residues: 27-334, D' 336-515, 'H' 517-554, 'L' 556-812 <SCH>
R;Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A;Title: Characterization of a complementary deoxyribonucleic acid coding for human an-
A;Reference number: I45961; MUID: 85023311; PMID: 6148961
A;Accession: I45961
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 706-743, 'R' 745-812 <MAL>
A;Cross-references: GB:K02935; NID: g163551; PID: g163552
R;Brunsholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Leigier, W.; Manneberg, M
Eur. J. Biochem. 114, 465-470, 1981
A;Title: Comparison of the primary structure of the N-terminal CNBr fragments of human
A;Reference number: S03735; MUID: 81212097; PMID: 7238497
A;Accession: S03736
A;Molecule type: protein
A;Residues: 27-83 <BRU>
C;Function:
A;Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a
ns the walls of the graafian follicle; also activates the urokinase-type plasminogen a
A;Pathway: fibrinolysis
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homolo-
C;Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasmin
F;1-26/Domain: signal sequence #status predicted <SIG>
F;8-103/Domain: plasminogen-related protein precursor homology <PLPH>
F;27-812/Product: plasminogen #status experimental <PRO>
F;27-103/Domain: activation peptide #status experimental <APT>
F;104-583,584-812/Product: plasmin #status experimental <MAT>
F;104-583/Domain: plasmin chain A #status experimental <ACH>
F;110-188/Domain: kringle homology <KR1>
F;192-269/Domain: kringle homology <KR2>
F;282-359/Domain: kringle homology <KR3>
F;384-461/Domain: kringle homology <KR4>
F;485-564/Domain: kringle homology <KR5>
F;584-812/Domain: plasmin chain B #status experimental <BCH>
F;584-805/Domain: trypsin homology <TRY>
F;56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,
bonds: #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;365/Binding site: carbohydrate (Ser) (covalent) #status experimental
F;624,667,762/Active site: His, Asp, Ser #status predicted

Query Match 19.7%; Score 156; DB 1; Length 812;
Best Local Similarity 28.3%; Pred. No. 1.7e-06;
Matches 47; Conservative 12; Mismatches 53; Indels 54; Gaps 10;
Qy 5 HQVPSNCDCLN-----GGTCVSNKYFSNIHW--CNCPKFKGGQHCIFDKSKT 49
Db 315 NRTPEPFCKNLEENYCRNPNNGEKAPWCYTTN--SEVRWEYCTIPS-----CSSLST 366
Qy 50 -----CYEGNGHFYRGKASTDTMGRCPLPWN SATVQLQTYHHAH---RS 89
Db 367 ERMDVPVPPEQTVPVQDCYHNGQSYRGTSSTITGRKQSWSS-----MTPHRLKTPTE 421
Qy 90 DALQLGLGKKNYCRNPNRRRPPWCYVQVGLKPLV--QECMVHDCAD 133
Db 422 NYPNAGL-TWNYCRNPDADKSPWCYT---TDPVRVWEFCNLKKCSE 463

Search completed: December 3, 2003, 14:44:15
Job time : 6.98276 secs

Qy 42 C-----EIDSKTCYEGNGHFYRGKASTDTMGRCPLPWN SATVQLQTYHHAHRS DAL 92
Db 296 CIRIGIPMADPINKNHKCYNSTGVDYRGTVSVTKSGRQCPWNS-----QYPHTHTPTAL 350
Qy 93 QLG--LGKHNKYCRNPNRRR--PWCY 114
Db 351 RFPELNGGHYCRNPGNQKEAPWCF 375

RESULT 29
B61545
plasmin (EC 3.4.21.7) precursor - sheep (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 17-Mar-1999
C;Accession: B61545; S28200
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID: 8905015; PMID: 3168975
A;Accession: B61545
A;Molecule type: protein
A;Residues: 1-37,38-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 5, 21-25, 1992
A;Title: Complete amino acid sequence of ovine miniplasminogen.
A;Reference number: S28200; MUID: 93149995; PMID: 1492092
A;Accession: S28200
A;Molecule type: protein
A;Residues: 118-460 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasmin; serine proteinase; z
F;1-37,38-117,118-460/Product: plasminogen (fragments) #status experimental <PRO>
F;1-37/Domain: activation peptide (fragment) #status experimental <APT>
F;38-117,118-230,231-460/Product: plasmin (fragments) #status experimental <MAT>
F;41-118/Domain: kringle homology <KR4>
F;118-460/Product: miniplasminogen #status experimental <MIN>
F;132-211/Domain: kringle homology <KR5>
F;226-460/Domain: plasmin chain B #status experimental <BCH>
F;231-453/Domain: trypsin homology <TRY>
F;272,315,410/Active site: His, Asp, Ser #status predicted

Query Match 19.9%; Score 157.5; DB 2; Length 460;
Best Local Similarity 38.2%; Pred. No. 7.5e-07;
Matches 34; Conservative 13; Mismatches 27; Indels 15; Gaps 4;
Qy 35 KKFGGQ-----HCEIDKSKTCYEGNGHFYRGKASTDTMGRCPLPWN SATV--LQOITYH 85
Db 20 KKLGRSVEDCAAKCE--EAGDCYHNGQGYRGTSSTITGRKQSWSSMIFPHRHQKTPTE 78
Qy 86 AHRSDALQLGLGKKNYCRNPNRRRPPWCY 114
Db 79 SYPNAGLTW-----NYCRNPDADKSPWCY 102

RESULT 30
PLBO
plasmin (EC 3.4.21.7) precursor - bovine
N;Alternate names: plasminogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C;Accession: S45046; A25835; I45961; S03736
R;Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A;Description: Cloning and characterization of the bovine plasminogen cDNA.
A;Reference number: S45046
A;Accession: S45046
A;Molecule type: mRNA
A;Residues: 1-812 <BER>
A;Cross-references: EMBL:X79402; NID: g494962; PID: CAA555939.1; PID: g494963
A;Experimental source: liver

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:53 ; Search time 4.33777 Seconds
(without alignments)
1463.563 Million cell updates/sec

Title: US-09-880-503-4
Perfect score: 793
Sequence: 1 SNELHQVPSNCDCLNGTGV.....QVGLKPLVQBMVHDCADGK 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	ID	Description
1	793	100.0	431	1 UROK HUMAN	P00749 homo sapien
2	713.5	90.0	433	1 UROK PAPCY	P16227 papio cynoc
3	608	76.7	442	1 UROK_PIG	P04185 sus scrofa
4	579	73.0	433	1 UROK_BOVIN	Q05589 bos taurus
5	571	72.0	432	1 UROK RAT	P29598 rattus norv
6	548	69.1	433	1 UROK_MOUSE	P06869 mus musculu
7	335.5	42.3	477	1 URT2_DESRO	P15638 desmodus ro
8	334.5	42.2	434	1 URT2_CHICK	P15120 gallus gall
9	328.5	41.4	431	1 URTB_DESRO	P08121 desmodus ro
10	328.5	41.4	562	1 TPA_HUMAN	P00750 homo sapien
11	321.5	40.5	559	1 TPA_RAT	P19637 rattus norv
12	315.5	39.8	559	1 TPA_MOUSE	P11214 mus musculu
13	310.5	39.2	477	1 URT1_DESRO	P08119 desmodus ro
14	300.5	37.9	566	1 TPA_BOVIN	Q04198 bos taurus
15	280.5	32.8	655	1 HGFA_HUMAN	Q04756 homo sapien
16	257	32.4	603	1 FA12_CAVPO	Q04962 cavia porce
17	250.5	31.6	653	1 HGFA_MOUSE	Q09098 mus musculu
18	233	29.4	615	1 FA12_HUMAN	P00748 homo sapien
19	223	28.1	394	1 URTG_DESRO	P49150 desmodus ro
20	217	27.4	593	1 FA12_BOVIN	P98140 bos taurus
21	169	21.3	685	1 ROR1_DROME	Q24488 dtrosophila
22	162.5	20.5	4548	1 APOA_HUMAN	P08519 homo sapien
23	161	20.3	473	1 KRM1_RAT	Q92484 rattus norv
24	161	20.3	1420	1 APOA_MACMU	P14417 macaca mula
25	159	20.1	452	1 KRM1_XENLA	Q90y90 xenopus lae
26	159	20.1	473	1 KRM1_MOUSE	Q99n43 mus musculu
27	158.5	20.0	937	1 ROR1_HUMAN	Q01973 homo sapien
28	158.5	20.0	937	1 ROR1_MOUSE	Q92139 mus musculu
29	156	19.7	475	1 KRM1_HUMAN	Q96m08 homo sapien
30	156	19.7	812	1 PLMN_BOVIN	P06868 bos taurus
31	155	19.5	790	1 PLMN_PIG	P06867 sus scrofa
32	154	19.4	810	1 PLMN_HUMAN	P00747 homo sapien
33	153.5	19.4	810	1 PLMN_ERIEU	Q29485 erinaceus e

RESULT 1

ID	UROK_HUMAN	STANDARD	PRT	431 AA
AC	P00749; Q15844; Q16618; Q969W6;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	20-MAR-1987 (Rel. 04, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)			
DE	(U-plasminogen activator).			
GN	PLAU.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85215647; PubMed=2987867;			
RA	Riccio A., Grimaldi G., Verde P., Sebastio G., Boast S., Blasi F.;			
RT	"The human urokinase-plasminogen activator gene and its promoter.;"			
RL	Nucleic Acids Res. 13:2759-2771(1985).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Holmes W.E., Pennica D., Blaber M., Rey M.W., Guenzler W.A.,			
RT	Steffens G.J., Heyneker H.L.;			
RL	"Cloning and expression of the gene for pro-urokinase in Escherichia coli.;"			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=86056954; PubMed=2415429;			
RA	Nagai M., Hitamatsu R., Kaneda T., Hayasuke N., Atimura H.,			
RT	Nishida M., Suyama T.;			
RL	"Molecular cloning of cDNA coding for human prepro-urokinase.;"			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85203359; PubMed=3888571;			
RA	Jacobs P., Cravador A., Loriau R., Brockly F., Colau B., Chuchana P.,			
RT	van Elsen A., Herzog A., Bollen A.;			
RL	"Molecular cloning, sequencing, and expression in Escherichia coli of human prepro-urokinase cDNA.;"			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,			
RT	Nickerson D.A.;			
RN	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.			
RN	[6]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lung;			
RX	MEDLINE=22386257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			

ALIGNMENTS

34	152.5	19.2	810	1	PLMN_MACMU	P12545 macaca mula
35	152	19.2	189	1	PLMN_RAT	Q01177 rattus norv
36	148.5	18.7	462	1	KRM2_HUMAN	Q0ncw0 homo sapien
37	148	18.7	943	1	ROR2_HUMAN	Q01974 homo sapien
38	148	18.7	944	1	ROR2_MOUSE	Q92138 mus musculu
39	147	18.5	812	1	PLMN_MOUSE	P20918 mus musculu
40	146.5	18.5	461	1	KRM2_MOUSE	Q8k1s7 mus musculu
41	144	18.2	711	1	HGFL_HUMAN	P26927 homo sapien
42	141.5	17.8	728	1	HGF_HUMAN	P14210 homo sapien
43	141	17.8	333	1	PLMN_CANFA	P80009 canis famil
44	139	17.5	716	1	HGFL_MOUSE	P26928 mus musculu
45	139	17.5	728	1	HGF_MOUSE	Q08048 mus musculu

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettanan M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE OF 66-431 FROM N.A.
 RX MEDLINE=84272706; PubMed=6589620;
 RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
 RT "Identification and primary sequence of an unspliced human urokinase
 RT poly(A)+ RNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
 RN [8]
 RP SEQUENCE OF 21-177.
 RX MEDLINE=83055084; PubMed=6754569;
 RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
 RA Flohe L.;
 RT "The primary structure of high molecular mass urokinase from human
 RT urine. The complete amino acid sequence of the A chain.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
 RN [9]
 RP SEQUENCE OF 156-176 AND 179-224.
 RX MEDLINE=83003608; PubMed=6749491;
 RA Schaller J., Nick H., Rickli E.E., Gillesen D., Lergier W.,
 RA Studer R.O.;
 RT "Human low-molecular-weight urinary urokinase. Partial
 RT characterization and preliminary sequence data of the two polypeptide
 RT chains.";
 RL Eur. J. Biochem. 125:251-257(1982).
 RN [10]
 RP SEQUENCE OF 158-410.
 RX MEDLINE=83055099; PubMed=6754572;
 RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
 RT "The complete amino acid sequence of low molecular mass urokinase
 RT from human urine.";
 RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RX MEDLINE=96000858; PubMed=8591045;
 RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
 RA Dobson C.M., Stuart D.I., Jones E.V.;
 RT "The crystal structure of the catalytic domain of human
 RT urokinase-type plasminogen activator.";
 RL Structure 3:681-691(1995).
 RN [12]
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
 RX MEDLINE=20266327; PubMed=10805774;
 RA Spel S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
 RA Bode W., Magdolen V., Huber R., Moroder L.;
 RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
 RT selective inhibitors of human urokinase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
 RN [13]
 RP STRUCTURE BY NMR.
 RX MEDLINE=89127526; PubMed=2536903;
 RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
 RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
 RT dimensional NMR.";
 RL Nature 337:579-582(1989).
 RN [14]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=93003110; PubMed=1377118;
 RA Li X., Smith R.A.G., Dobson C.M.;

RT "Sequential 1H NMR assignments and secondary structure of the kringle
 RT domain from urokinase.";
 RL Biochemistry 31:9562-9571(1992).
 RN [15]
 RP STRUCTURE BY NMR OF 67-155.
 RX MEDLINE=94149701; PubMed=8107091;
 RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
 RT "Solution structure of the kringle domain from urokinase-type
 RT plasminogen activator.";
 RL J. Mol. Biol. 235:1548-1559(1994).
 RN [16]
 RP VARIANT LEU-141.
 RX MEDLINE=96186279; PubMed=8652631;
 RA Yoshimoto M., Ushiyama Y., Sakai M., Tamaki S., Hara H., Takahashi K.,
 RA Sawasaki Y., Hanada K.;
 RT "Characterization of single chain urokinase-type plasminogen
 RT activator with a novel amino-acid substitution in the kringle
 RT structure.";
 RL Biochim. Biophys. Acta 1293:83-89(1996).
 RN [17]
 RP VARIANT LEU-141.
 RX MEDLINE=97218551; PubMed=9065988;
 RA Conne B., Berczy M., Belin D.;
 RT "Detection of polymorphisms in the human urokinase-type plasminogen
 RT activator gene.";
 RL Thromb. Haemost. 77:434-435(1997).
 RN [18]
 RP ERRATUM.
 RA Conne B., Berczy M., Belin D.;
 RL Thromb. Haemost. 78:973-973(1997).
 RN [19]
 RP VARIANT LEU-141.
 RX MEDLINE=97337920; PubMed=9194591;
 RA Turkmen B., Schmitt M., Schmalfeldt B., Trommler P., Hell W.,
 RA Creutzburg S., Graeff H., Magdolen V.;
 RT "Mutational analysis of the genes encoding urokinase-type plasminogen
 RT activator (upa) and its inhibitor PAI-1 in advanced ovarian cancer.";
 RL Electrophoresis 18:686-689(1997).
 CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
 CC THERAPY OF THROMBOLYTIC DISORDERS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 CC plasminogen to form plasmin.
 CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
 CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
 CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
 CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
 CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
 CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02419; CAA26268.1; -
 DR EMBL; M15476; AA61253.1; -
 DR EMBL; D00244; BAA00175.1; -
 DR EMBL; D11143; BAA01919.1; -
 DR EMBL; X02760; CAA26535.1; -
 DR EMBL; AF377330; AAK53822.1; -
 DR EMBL; BC013575; AAH13575.1; -
 DR EMBL; K03226; AAC97138.1; -
 DR EMBL; K02286; AA61452.1; -
 DR EMBL; A21571; CAA01559.1; -
 DR EMBL; A18397; CAA01390.1; -
 DR PIR; A00931; UKHU.
 DR PDB; 1KDU; 31-OCT-93.

```
Query Match      100.0%; Score 793; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.6e-70;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 80
QY 61 KASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 81 KASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 2
UROK_PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR ENBL; X51935; CAA36200.1; -.
DR PIR; S14687; URBAY.
DR HSSP; P00749; ILMW.
DR MEROPS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
```

```
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match      90.0%; Score 713.5; DB 1; Length 433;
Best Local Similarity 91.9%; Pred. No. 1.5e-62;
Matches 124; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 1 SNEHQVPSNCDCLNGGTCVSNKYFNSIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 21 SREL-QVPSDCGLNGGTCVSNKYFNSIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 79
QY 61 KASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 120
DB 80 KASTDTMGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVGLK 139
QY 121 PLVQECMVHDCADGK 135
DB 140 QRVQECMVHDCADGK 154

RESULT 3
UROK_PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator.";
```

Nucleic Acids Res. 12:9525-9541 (1984).

[2]
RN REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: Contains 1 kringle domain.
CC -|- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01648; CAA25806.1; -;
DR EMBL; X02724; CAA26511.1; -;
DR PIR; A00932; UKFG.
DR HSP; P00749; 1KDU.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE_NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Plasminogen activation; Hydrolase;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442
FT CHAIN 21 188
FT CHAIN 190 442
FT DOMAIN 29 65
FT DOMAIN 72 153
FT DOMAIN 154 189
FT DOMAIN 190 442
FT CARBOHYD 152 152
FT DISULFID 33 41
FT DISULFID 35 53
FT DISULFID 55 64
FT DISULFID 179 310
FT DISULFID 220 236
FT DISULFID 228 299
FT DISULFID 324 393
FT DISULFID 356 372
FT DISULFID 383 411
FT ACT_SITE 235 235
FT ACT_SITE 286 286
FT ACT_SITE 387 387
FT CONFLICT 241 241
FT CONFLICT 242 242
FT CONFLICT 288 288
SQ SEQUENCE 442 AA; 49116 MW; E832FCF501321EE CRC64;

Query Match 76.7%; Score 608; DB 1; Length 442;
Best Local Similarity 78.1%; Pred. No. 3.2e-52;
Matches 107; Conservative 11; Mismatches 17; Indels 2; Gaps 1;

QY 1 SNELHQV--PSNCDCLNGGTCVSNKYFSNIHWCNPKKFGGQHCETDKSKTCYEGNGHYF 58
|:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 SHELHQESGASNGCLNGGKCVSYKYFSNIQRCSPKFKQGEHCETDTSQTCFEGNGHSY 80
|:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 59 RGKASTDTWGRPCLPWNSATVLTQOTYAHRESDALQLGLGKHNYCRNPDNRPRPCYVQVG 118
|:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 RGRANTINTGRPCLPWNSATVLTNTTHAHRPDALQLGLGKHNYCRNPDNRPRPCYVQVG 140
|:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 119 LKPLVQECMVHDCADGK 135
|:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 LKQLVQECMVPCSGGE 157
|:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
UROK_BOVIN UROK_BOVIN STANDARD; PRT; 433 AA.
ID UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (BC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schleuning W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
RL Gene 125:177-183(1993).
RN [2]
RP SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA";
RL int. Dairy J. 5:605-617(1995).
CC -|- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -|- INDUCTION: By retinoic acid.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: Contains 1 kringle domain.
CC -|- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L03546; AAA51419.1; -;
DR EMBL; X85801; CAA59796.1; -;
DR PIR; JN0560; JN0560.
DR HSP; P00749; 1LMW.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.


```

FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match
Best Local Similarity 72.0%; Score 571; DB 1; Length 432;
Matches 98; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 9 SNCDCLNGTGVSNKYFNSIHWCNCPKPGGQHCEIDKSKTCYEGNGHYPYRGKASTDTMG 68
DB 29 SNGCGQGGVGVSVKYFSSIRRCSPKPKFGHCEIDKSKTCYHGNGGSGYRGKANTDTKG 88
QY 69 RPLCPWNSATVLQOQTYHAHRSDALQLGLGKHNVCNPNRRRPWCYVGVGLKPLVQECMV 128
DB 89 RPLCLAWSFAVLQOQTYHAHRSDALSLGLGKHNVCNPNRRRPWCYVGVGLKPLVQECMV 148
QY 129 HDCADGK 135
DB 149 QDCSLSK 155

RESULT 6
UROK_MOUSE STANDARD; PRT; 433 AA.
ID AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=85179474; PubMed=2985383;
RX Belin D., Vassalli J.-D., Combepine C., Godeau F., Nagamine Y.,
RA Reich E., Koehler H.P., Duvoisin R.M.;
RA "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RT Eur. J. Biochem. 148:225-232(1985).
RN [2]
SEQUENCE FROM N.A.
RP MEDLINE=88163489; PubMed=2831940;
RX Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279(1987).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X02389; CAA26231.1; -
CC EMBL; M17922; AAA40539.1; -
CC PIR; A29420; UKWS
CC HSSP; P00749; 1KDU.
CC MEROPS; S01.231; -.

```


OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
ON NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidel W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Salivary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
RA Register R.B., Jacobs J.W., Dixon R.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
RT salivary plasminogen activator.";
RL J. Biol. Chem. 264:17947-17952(1989).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidel W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
CC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
CC THE PRESENCE OF FIBRIN I.
CC -!- SUBUNIT: Monomer.
CC -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
CC AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
CC STIMULATION OF ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; MG3988; AAA31593.1; -;
DR EMBL; J05082; AAA31596.1; -;
DR PIR; JS0598; JS0598.
DR HSP; P98119; I451.
DR MEROPS; S01.232; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibinctn1.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fni; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation: Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Multigene family.
FT SIGNAL 1 36
FT CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 2.
FT DOMAIN 40 82 FIBRONECTIN TYPE-I.
FT DOMAIN 83 121 EGF-LIKE.
FT DOMAIN 128 209 KRINGLE.
FT DOMAIN 225 477 SERINE PROTEASE.
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 321 321 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 428 428 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 42 72 BY SIMILARITY.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 87 98 BY SIMILARITY.
FT DISULFID 92 109 BY SIMILARITY.
FT DISULFID 111 120 BY SIMILARITY.
FT DISULFID 128 209 BY SIMILARITY.
FT DISULFID 149 191 BY SIMILARITY.
FT DISULFID 180 204 BY SIMILARITY.
FT DISULFID 214 345 BY SIMILARITY.
FT DISULFID 257 273 BY SIMILARITY.
FT DISULFID 265 334 BY SIMILARITY.
FT DISULFID 359 434 BY SIMILARITY.
FT DISULFID 391 407 BY SIMILARITY.
FT DISULFID 424 452 BY SIMILARITY.
FT CARBOHYD 185 185 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC...) (POTENTIAL).
FT CONFLICT 403 403 N -> K (IN REF. 2).
FT CONFLICT 417 417 Y -> H (IN REF. 2).
FT CONFLICT 435 435 M -> R (IN REF. 2).
SQ SEQUENCE 477 AA; 53719 MW; 17486555C0S05077C CRC64;

Query Match 42.3%; Score 335.5; DB 1; Length 477;
Best Local Similarity 46.3%; Pred. No. 1.4e-25;
Matches 62; Conservative 17; Mismatches 50; Indels 5; Gaps 2;

QY 3 ELHQP-----SNCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHFY 58
Db 78 QCHTVPKSCSELRCFNGGTCWQAASDF-VCCPFGYTGKQCEVDTHATCYRDQGVY 136

QY 59 RGKASTDTWGRPLCPWNSATVLQOITYAHRSDALQLGLGKHNYCRNPDNRERPCVQVG 118
Db 137 RGTWTSSEGAQCIWNNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDNNSKPCVYVKA 196

QY 119 LKPLVQECMVHDC 132
Db 197 SKFILEFCVPVCS 210

RESULT 8
UROK_CHICK STANDARD; PRT; 434 AA.
AC F15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [9]
RP SEQUENCE OF 212-361 FROM N.A.
RX MEDLINE=8316956; PubMed=6572897;
RA Edlund T., Ny T., Raanby M., Heden L.-O., Palm G., Holmgren E.,
RA Josephson S.;
RT "Isolation of cDNA sequences coding for a part of human tissue
RT plasminogen activator.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:349-352(1983).
RN [10]
RP SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=85289338; PubMed=3161893;
RA Fisher R., Waller E.K., Grossi G., Thompson D., Tizard R.,
RA Schleuning W.-D.;
RT "Isolation and characterization of the human tissue-type plasminogen
RT activator structural gene including its 5' flanking region.";
RL J. Biol. Chem. 260:11223-11230(1985).
RN [11]
RP SEQUENCE OF 31-562 FROM N.A.
RX MEDLINE=91291340; PubMed=1368681;
RA Itagaki Y., Yasuda H., Morinaga T., Mitsuda S., Higashio K.;
RT "Purification and characterization of tissue plasminogen activator
RT secreted by human embryonic lung diploid fibroblasts, IMR-90 cells.";
RL Agric. Biol. Chem. 55:1225-1232(1991).
RN [12]
RP SEQUENCE OF 36-562.
RX TISSUE-Melanoma;
RC MEDLINE=8500468; PubMed=6433976;
RA Pohl G., Kaelstroem M., Bergsdorf N., Wallen P., Joernvall H.;
RT "Tissue plasminogen activator: peptide analyses confirm an indirectly
RT derived amino acid sequence, identify the active site serine residue,
RT establish glycosylation sites, and localize variant differences.";
RL Biochemistry 23:3701-3707(1984).
RN [13]
RP SEQUENCE OF 33-52 AND 311-330.
RX TISSUE-Melanoma;
RC MEDLINE=83209620; PubMed=6682760;
RA Wallen P., Pohl G., Bergsdorf N., Raanby M., Ny T., Joernvall H.;
RT "Purification and characterization of a melanoma cell plasminogen
RT activator.";
RL Eur. J. Biochem. 132:681-686(1983).
RN [14]
RP STRUCTURE OF CARBOHYDRATES.
RX MEDLINE=90092112; PubMed=2513186;
RA Pfeiffer G., Schmidt M., Strube K.-H., Geyer R.;
RT "Carbohydrate structure of recombinant human uterine tissue
RT plasminogen activator expressed in mouse epithelial cells.";
RL Eur. J. Biochem. 186:273-286(1989).
RN [15]
RP CARBOHYDRATE-LINKAGE SITE THR-96.
RX MEDLINE=91159408; PubMed=1900431;
RA Harris R.J., Leonard C.K., Guzzetta A.W., Spellman M.W.;
RT "Tissue plasminogen activator has an O-linked fucose attached to
RT threonine-61 in the epidermal growth factor domain.";
RL Biochemistry 30:2311-2314(1991).
RN [16]
RP DISULFIDE BONDS IN KRINGLE 2.
RX MEDLINE=91244765; PubMed=1645336;
RA Vlahos C.J., Wilhelm O.G., Hassell T., Jaskunas S.R., Bang N.U.;
RT "Disulfide pairing of the recombinant kringle-2 domain of tissue
RT plasminogen activator produced in Escherichia coli.";
RL J. Biol. Chem. 266:10070-10072(1991).
RN [17]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=96200985; PubMed=8613982;
RA Lamba D., Bauer M., Huber R., Fischer S., Rudolph R., Kohnert U.,
RA Bode W.;
RT "The 2.3 A crystal structure of the catalytic domain of recombinant
RT two-chain human tissue-type plasminogen activator.";

RL J. Mol. Biol. 258:117-135(1996).
RN [18]
RP X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF CATALYTIC DOMAIN.
RX MEDLINE=97449126; PubMed=9305622;
RA Renatus M., Engh R.A., Stubbs M.T., Huber R., Fischer S., Kohnert U.,
RA Bode W.;
RT "Lysine 156 promotes the anomalous proenzyme activity of tPA: X-ray
RT crystal structure of single-chain human tPA.";
RL EMBO J. 16:4797-4805(1997).
RN [19]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF KRINGLE 2.
RX MEDLINE=92118803; PubMed=1310033;
RA de Vos A., Ultsch M.H., Kelley R.F., Padmanabhan K., Tulinsky A.,
RA Westbrook M.L., Kossiakof A.A.;
RT "Crystal structure of the kringle 2 domain of tissue plasminogen
RT activator at 2.4-A resolution.";
RL Biochemistry 31:270-279(1992).
RN [20]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=90122799; PubMed=2558718;
RA Byeon I.-J.L., Kelley R.F., Ilinas M.;
RT "1H NMR structural characterization of a recombinant kringle 2 domain
RT from human tissue-type plasminogen activator.";
RL Biochemistry 28:9350-9360(1989).
RN [21]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=91200042; PubMed=1901789;
RA Byeon I.-J.L., Kelley R.F., Ilinas M.;
RT "Kringle-2 domain of the tissue-type plasminogen activator. 1H-NMR
RT assignments and secondary structure.";
RL Eur. J. Biochem. 197:155-165(1991).
RN [22]
RP STRUCTURE BY NMR OF KRINGLE 2.
RX MEDLINE=92106329; PubMed=1762144;
RA Byeon I.-J.L., Ilinas M.;
RT "Solution structure of the tissue-type plasminogen activator kringle
RT 2 domain complexed to 6-aminohexanoic acid an antifibrinolytic
RT drug.";
RL J. Mol. Biol. 222:1035-1051(1991).
RN [23]
Query Match 41.4%; Score 328.5; DB 1; Length 562;
Best Local Similarity 46.3%; Pred. No. 8.1e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;
Qy 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNPKKFGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSEPRCFNGGTQQALYFSDF-VCQCPGEGFAGKCEIDTRATCYEDQGTSY 135
Qy 59 RGKASTDTMGPCPLPWN SATVLQOTYHAHRS DALQLGKHN YCNPENRRRPHCYQVG 118
Db 136 RGTWSTAESGAECTNWNSSALAKPKYSGRRPDALRLGLGNHNYCRNPDRDSKWCYVPKA 195
Qy 119 LKPLVQECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211
RESULT 11
TPA_RAT STANDARD; PRT; 559 AA.
ID TPA_RAT
AC P19637;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator).
GN PLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]


```

FT DISULFID 108 133 BY SIMILARITY.
FT DISULFID 122 148 BY SIMILARITY.
FT DISULFID 164 175 BY SIMILARITY.
FT DISULFID 169 186 BY SIMILARITY.
FT DISULFID 188 197 BY SIMILARITY.
FT DISULFID 202 230 BY SIMILARITY.
FT DISULFID 228 237 BY SIMILARITY.
FT DISULFID 245 256 BY SIMILARITY.
FT DISULFID 250 267 BY SIMILARITY.
FT DISULFID 269 278 BY SIMILARITY.
FT DISULFID 286 367 BY SIMILARITY.
FT DISULFID 307 349 BY SIMILARITY.
FT DISULFID 338 362 BY SIMILARITY.
FT DISULFID 394 521 INTERCHAIN (BY SIMILARITY).
FT DISULFID 432 448 BY SIMILARITY.
FT DISULFID 440 510 BY SIMILARITY.
FT DISULFID 535 604 BY SIMILARITY.
FT DISULFID 567 583 BY SIMILARITY.
FT DISULFID 594 622 BY SIMILARITY.
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 644 644 R -> Q (IN REF. 2).
SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1B1B862ED7 CRC64;

Query Match 32.8%; Score 260.5; DB 1; Length 655;
Best Local Similarity 43.7%; Pred. No. 4.1e-18;
Matches 52; Conservative 8; Mismatches 42; Indels 17; Gaps 2;

QY 5 HQVPSNCDLNGTGVSNKYFSNIHW-----CNCPKFGGHCIDKSKTCYSGNGH 56
Db 242 HTACLSPLNGTGC-----HLVATGTVACPPGAGFLCNIDEDRCFLNGT 292

QY 57 FYRGKASTDTMGPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRPPCYV 115
Db 293 GYRGVASTSASGLSCLAWNSDLLYQLHVDVSGAALLGLGPHAYCRNPDRRPPCYV 351

RESULT 16
FA12_CAVPO STANDARD; PRT; 603 AA.
AC Q04962;
DC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxId=10141;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RX MEDLINE=93003367; PubMed=1390917;
RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA Kambara T., Okabe H.;
RT "Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule.";
RL Biochim. Biophys. Acta 1159:113-121(1992).
CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC CATALYTIC ACTIVITY: Cleaves selectively Arg-|-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIIA AND THEN TO BETA-FACTOR XIIIA. ALPHA-FACTOR
CC XIIIA ACTIVATES FACTOR XI TO FACTOR XIA.

```

```

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68615; CAA48600.1; -.
DR PIR; S28941; S28941.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON TER 1 1
FT SIGNAL <1 18
FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
FT DOMAIN 93 130 EGF-LIKE 1.
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
FT DOMAIN 173 209 EGF-LIKE 2.
FT DOMAIN 216 294 KRINGLE.
FT DOMAIN 312 342 PRO-RICH.
FT DOMAIN 359 603 SERINE PROTEASE.
FT ACT SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 118 BY SIMILARITY.
FT DISULFID 120 129 BY SIMILARITY.
FT DISULFID 134 162 BY SIMILARITY.
FT DISULFID 160 169 BY SIMILARITY.
FT DISULFID 177 188 BY SIMILARITY.

```

```

FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.
FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9ED59 CRC64;

Query Match 32.4%; Score 257; DB 1; Length 603;
Best Local Similarity 40.8%; Pred. No. 8.4e-18;
Matches 49; Conservative 17; Mismatches 46; Indels 8; Gaps 3;

QY 13 CLNGGTCVNKFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
Db 182 CLNGGRCLE--VEGHLLDCPCMGYTGPFCDLDTTASCYEGRGVSYRGMARITVSGAKCQ 238
QY 73 PMSATVLQOTVHAHRS-D-ALQILGLCKHNYCRPNDRRPPWCYVQVGLKPLVQECMVHDC 131
Db 239 RWAS-----EATRYNMTAEQALRGHLGHHTFCRPNDRTPWPCFVWMGNRLSMEYCDLAQC 294

RESULT 17
HGFA_MOUSE
ID HGFA_MOUSE STANDARD; PRT; 653 AA.
AC QPR098; Q9UKV4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RN SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RA Itoh H., Kataoka H., Koono H.;
RT "Mouse hepatocyte growth factor activator.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA van Adelsberg J.S., Sehgal S., Kukes A., Brady C., Barasch J.,
RA Yang J., Huan Y.;
RT "Activation of HGF by endogenous HGF activator is required for
RT metanephric kidney morphogenesis in vitro.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY CONVERTING
CC IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF099017; AAF02489.1; -.
DR EMBL; AF224724; AAF34712.1; -.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.228; -.
DR MGD; MGI:1859281; Hgfac.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF 2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibnctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; ENTPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KE; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02400; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_SER; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
KW EGF-like domain; Repeat; Zymogen.
FT SIGNAL 1 29 BY SIMILARITY.
FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT
FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG
FT CHAIN 105 145 CHAIN
FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
FT DOMAIN 157 195 EGF-LIKE 1.
FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
FT DOMAIN 238 276 EGF-LIKE 2.
FT DOMAIN 283 364 KRINGLE.
FT DOMAIN 406 653 SERINE PROTEASE.
FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 105 130 BY SIMILARITY.
FT DISULFID 119 145 BY SIMILARITY.
FT DISULFID 161 172 BY SIMILARITY.
FT DISULFID 166 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 199 227 BY SIMILARITY.
FT DISULFID 225 234 BY SIMILARITY.
FT DISULFID 242 253 BY SIMILARITY.
FT DISULFID 247 264 BY SIMILARITY.
FT DISULFID 266 275 BY SIMILARITY.
FT DISULFID 283 364 BY SIMILARITY.
FT DISULFID 304 346 BY SIMILARITY.
FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
FT DISULFID 430 446 BY SIMILARITY.

```

```
FT DISULFID 438 508 BY SIMILARITY.
FT DISULFID 533 602 BY SIMILARITY.
FT DISULFID 565 581 BY SIMILARITY.
FT DISULFID 592 620 BY SIMILARITY.
FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 164 164 G -> W (IN REF. 2).
SQ SEQUENCE 653 AA; 70567 MW; 88B4B20255DF7FDC CRC64;

Query Match 31.6%; Score 250.5; DB 1; Length 653;
Best Local Similarity 42.0%; Pred. No. 3.9e-17;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

QY 5 HQVPSKCDLNGTGVSNKFSNIHW-----CNCPKFGQGHCEIDKSKTCYEGNGH 56
DB 239 HTACLSPCLNGTGC-----HLIVGTGTVCTCPLGYAGRFNCNIVPTERCFLNGT 289
QY 57 FYRGKASTDTMGRCPLPWNBSATVLQOYTHAHRSDALQLGLGKHCNCRPNRRRPPWCYV 115
DB 290 EYRGVASTASGLSCLAWNSDLIYQELHVDVAAVLLGLGPHAYCENPKDKRPPWCYV 348

RESULT 18
FA12_HUMAN
ID FA12_HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF).
GN FI2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88007593; PubMed=2888762;
RA Cool D.E., McGilivray R.T.A.;
RT "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region.";
RL J. Biol. Chem. 262:13662-13673(1987).
RN [2]
RP SEQUENCE FROM N.A. AND VARIANTS ALA-207; ASP-545 AND HIS-605.
RA Rieder M.J., Armel T.Z., Carrington D.P., Ozuna M., Kuldanek S.A.,
RA Rajkumar N., Toth E.J., Yi Q., Nickerson D.A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RL Nucleic Acids Res. 14:3146-3146(1986).
RN [4]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGilivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor XIIa.";
RL J. Biol. Chem. 260:13666-13676(1985).
RN [5]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
```

```
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RL Biochemistry 25:1525-1528(1986).
RN [6]
RP SEQUENCE OF 20-379.
RX MEDLINE=85182674; PubMed=3886654;
RA McMullen B.A., Fujikawa K.;
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa
RT (activated Hageman factor).";
RL J. Biol. Chem. 260:5328-5341(1985).
RN [7]
RP SEQUENCE OF 354-362 AND 373-615.
RX MEDLINE=83291041; PubMed=6604055;
RA Fujikawa K., McMullen B.A.;
RT "Amino acid sequence of human beta-factor XIIa.";
RL J. Biol. Chem. 258:10924-10933(1983).
RN [8]
RP SEQUENCE OF 561-615 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=96133302; PubMed=8528215;
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;
RT "The novel acceptor splice site mutation 11396(G->A) in the factor
RT XII gene causes a truncated transcript in cross-reacting material
RT negative patients.";
RL Hum. Mol. Genet. 4:1235-1237(1995).
RN [9]
RP CARBOHYDRATE-LINKAGE SITE THR-109.
RX MEDLINE=92184750; PubMed=1544894;
RA Harris R.J., Ling V.T., Spellman M.W.;
RT "O-linked fucose is present in the first epidermal growth factor
RT domain of factor XII but not protein C.";
RL J. Biol. Chem. 267:5102-5107(1992).
RN [10]
RP VARIANT WASHINGTON D.C. SER-590.
RX MEDLINE=90046788; PubMed=2510163;
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alving B.,
RA Saito H.;
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive
RT factor XIIa results from Cys-571->Ser substitution.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).
RN [11]
RP VARIANT LOCARNO PRO-372.
RX MEDLINE=94325559; PubMed=8049433;
RA Hovinga J.K., Schaller J., Stricker H., Willemin W.A., Furlan M.,
RA Laemmle B.;
RT "Coagulation factor XII Locarno: the functional defect is caused by
RT the amino acid substitution Arg-353->Pro leading to loss of a
RT kallikrein cleavage site.";
RL Blood 84:1173-1181(1994).
RN [12]
RP VARIANT TENRI CYS-53.
RX MEDLINE=99290785; PubMed=10361128;
RA Kondo S., Tokunaga F., Kawano S., Oono Y., Kumagai S., Koide T.;
RT "Factor XII Tenri, a novel cross-reacting material negative factor XII
RT deficiency, occurs through a protease-mediated degradation.";
RL Blood 93:4300-4308(1999).
CC -I- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -I- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-lle bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -I- PTM: O- AND N-GLYCOSYLATED.
CC -I- DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.
CC -I- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR
CC XIIa ACTIVATES FACTOR XI TO FACTOR XIa.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 2 EGF-like domains.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 1 fibronectin type II domain.
```

CC	-!- SIMILARITY: Contains 1 kringle domain.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M31315; AAA70225.1; --
DR	EMBL; AP538691; RAM97932.1; --
DR	EMBL; M11723; AAAS1986.1; --
DR	EMBL; M17466; AAB59490.1; --
DR	EMBL; M17464; AAB59490.1; JOINED.
DR	EMBL; M17465; AAB59490.1; JOINED.
DR	EMBL; M13147; AAA70224.1; --
DR	EMBL; U71274; AAB51203.1; --
DR	FIR; A29411; KFHU12.
DR	HSSP; P00763; LDPO.
DR	MEROPS; S01.211; --
DR	Genew; HGNC:3530; F12.
DR	MIM; 234000; --
DR	GO; GO:0003805; F.blood coagulation factor XI activity; TAS.
DR	GO; GO:0003806; F.blood coagulation factor XII activity; TAS.
DR	GO; GO:0007596; F.blood coagulation; TAS.
DR	InterPro; IPR001314; Chymotrypsin.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; Fibronctn1.
DR	InterPro; IPR000562; FN_Type_II.
DR	InterPro; IPR006210; IEFG.
DR	InterPro; IPR000001; Kringle.
DR	InterPro; IPR001254; Ser_protease_Try.
DR	Pfam; PF000008; EGF; 2.
DR	Pfam; PF000039; fn1; 1.
DR	Pfam; PF000040; fn2; 1.
DR	Pfam; PF000051; kringle1; 1.
DR	Pfam; PF000089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00013; FNTYPEII.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; PD000095; FN_Type_II; 1.
DR	ProDom; PD000395; kringle1; 1.
DR	SMART; SM00181; EGF; 2.
DR	SMART; SM00058; FN1; 1.
DR	SMART; SM00059; FN2; 1.
DR	SMART; SM00130; KR; 1.
DR	SMART; SM00020; TRYD_SPC; 1.
DR	PROSITE; PS00022; EGF_1; 2.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PS00021; KRINGLE_2; 1.
DR	PROSITE; PS00070; KRINGLE_DOM; 1.
DR	PROSITE; PS00240; TRYPsin_DOM; 1.
DR	PROSITE; PS00134; TRYPsin_HIS; 1.
DR	PROSITE; PS00135; TRYPsin_SER; 1.
KW	Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW	Hydroxylase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW	Polymorphism; Disease mutation.
FT	SIGNAL 1 19
FT	CHAIN 20 372
FT	CHAIN 373 615
FT	CHAIN 354 362
FT	BETA-FACTOR XIAA PART 1.
FT	CHAIN 373 615
FT	BETA-FACTOR XIAA PART 2.
FT	DOMAIN 47 88
FT	FIBRONECTIN TYPE-II.
FT	EGF-LIKE 1.
FT	DOMAIN 94 131
FT	FIBRONECTIN TYPE-I.
FT	DOMAIN 133 173
FT	EGF-LIKE 2.
FT	DOMAIN 174 210
FT	KRINGLE.
FT	DOMAIN 217 295
FT	PRO-RICH.
FT	SERINE PROTEASE.
FT	DOMAIN 373 615
FT	O-LINKED (FUC). 109
FT	CARBOHYD

FT	CARBOHYD	249	249	N-LINKED (GLCNAC...).
FT	CARBOHYD	299	299	O-LINKED (POTENTIAL).
FT	CARBOHYD	305	305	O-LINKED (POTENTIAL).
FT	CARBOHYD	308	308	O-LINKED (POTENTIAL).
FT	CARBOHYD	328	328	O-LINKED (POTENTIAL).
FT	CARBOHYD	329	329	O-LINKED (POTENTIAL).
FT	CARBOHYD	337	337	O-LINKED (POTENTIAL).
FT	ACT_SITE	412	412	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	461	461	CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match 29.4%; Score 233; DB 1; Length 615;
Best Local Similarity 42.3%; Pred.No.1.9e-15;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

QY	13	CLNGGTCVSNKYFSNIHWCNCPKPGGHCEIDSKSCYEGRGHFYRGKASTDTMGAPCL	72
DB	183	CLHGRCLE---VEGRLCHCPGVGTGFCDDVDTKASYCDRGSLYGLARTTLTGAPQC	239
QY	73	PWNSATVLQQTY-HAHRSDALQLGLGKHNYCRNDNRPPWCYV	115
DB	240	PWAS----EATRVNTAEQARNWLGLGHAFCRPNDFDPWCFV	279

RESULT 19
URTG DESRO STANDARD; PRT; 394 AA.
AC P49150;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP TISSUE-Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P., Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";
PL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T., Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W., Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
CC -!- SUBUNIT: Monomer.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; M63990; AAA31595.1; --

CC	-!- SIMILARITY: Contains 1 kringle domain.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; M31315; AAA70225.1; --
DR	EMBL; AP538691; RAM97932.1; --
DR	EMBL; M11723; AAAS1986.1; --
DR	EMBL; M17466; AAB59490.1; --
DR	EMBL; M17464; AAB59490.1; JOINED.
DR	EMBL; M17465; AAB59490.1; JOINED.
DR	EMBL; M13147; AAA70224.1; --
DR	EMBL; U71274; AAB51203.1; --
DR	FIR; A29411; KFHU12.
DR	HSSP; P00763; LDPO.
DR	MEROPS; S01.211; --
DR	Genew; HGNC:3530; F12.
DR	MIM; 234000; --
DR	GO; GO:0003805; F.blood coagulation factor XI activity; TAS.
DR	GO; GO:0003806; F.blood coagulation factor XII activity; TAS.
DR	GO; GO:0007596; F.blood coagulation; TAS.
DR	InterPro; IPR001314; Chymotrypsin.
DR	InterPro; IPR006209; EGF like.
DR	InterPro; IPR000083; Fibronctn1.
DR	InterPro; IPR000562; FN_Type_II.
DR	InterPro; IPR006210; IEFG.
DR	InterPro; IPR000001; Kringle.
DR	InterPro; IPR001254; Ser_protease_Try.
DR	Pfam; PF000008; EGF; 2.
DR	Pfam; PF000039; fn1; 1.
DR	Pfam; PF000040; fn2; 1.
DR	Pfam; PF000051; kringle1; 1.
DR	Pfam; PF000089; trypsin; 1.
DR	PRINTS; PR00722; CHYMOTRYPSIN.
DR	PRINTS; PR00013; FNTYPEII.
DR	PRINTS; PR00018; KRINGLE.
DR	ProDom; PD000095; FN_Type_II; 1.
DR	ProDom; PD000395; kringle1; 1.
DR	SMART; SM00181; EGF; 2.
DR	SMART; SM000058; FN1; 1.
DR	SMART; SM000059; FN2; 1.
DR	SMART; SM00130; KR; 1.
DR	SMART; SM000020; Tryd_SPC; 1.
DR	PROSITE; PS000022; EGF_1; 2.
DR	PROSITE; PS01186; EGF_2; 1.
DR	PROSITE; PS01253; FIBRONECTIN_1; 1.
DR	PROSITE; PS00023; FIBRONECTIN_2; 1.
DR	PROSITE; PS00021; KRINGLE_1; 1.
DR	PROSITE; PS00021; KRINGLE_2; 1.
DR	PROSITE; PS00070; KRINGLE_DOM; 1.
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
DR	PROSITE; PS00135; TRYPSIN_SER; 1.
KW	Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW	Hydroxylase; Fibrinolysis; Signal; EGF-like domain; Repeat; Zymogen;
KW	Polymorphism; Disease mutation.
FT	SIGNAL 1 19
FT	CHAIN 20 372
FT	CHAIN 373 615
FT	CHAIN 354 362
FT	BETA-FACTOR XIAA PART 1.
FT	CHAIN 373 615
FT	BETA-FACTOR XIAA PART 2.
FT	DOMAIN 47 88
FT	FIBRONECTIN TYPE-II.
FT	EGF-LIKE 1.
FT	DOMAIN 94 131
FT	FIBRONECTIN TYPE-I.
FT	DOMAIN 133 173
FT	EGF-LIKE 2.
FT	DOMAIN 174 210
FT	KRINGLE.
FT	DOMAIN 217 295
FT	PRO-RICH.
FT	SERINE PROTEASE.
FT	DOMAIN 373 615
FT	O-LINKED (FUC). 109
FT	CARBOHYD

FT	CARBOHYD	249	249	N-LINKED (GLCNAC...).
FT	CARBOHYD	299	299	O-LINKED (POTENTIAL).
FT	CARBOHYD	305	305	O-LINKED (POTENTIAL).
FT	CARBOHYD	308	308	O-LINKED (POTENTIAL).
FT	CARBOHYD	328	328	O-LINKED (POTENTIAL).
FT	CARBOHYD	329	329	O-LINKED (POTENTIAL).
FT	CARBOHYD	337	337	O-LINKED (POTENTIAL).
FT	ACT_SITE	412	412	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	461	461	CHARGE RELAY SYSTEM (BY SIMILARITY).

Query Match	29.4%; Score 233; DB 1; Length 615;
Best Local Similarity	42.3%; Pred.No.1.9e-15;
Matches	44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

QY	13 CLNGGTCVSNKFSNIHWCNCPKPGGHCEIDSKSCYEGRGHFYRGKASTDTMGAPCL 72
DB	183 CLHGRCLE---VEGRLCHCPGVGTGFCDDVDTKASYCDRGSLSYGLARTTLTGAPQC 239
QY	73 PNWSATVLQQTY-HAHRSDALQLGLGKHNYCRNPDRRRPWCYV 115
DB	240 PWAS----EATRVNTAEQAQNWLGGHAFGRPNDFNDIRPWCFV 279

RESULT 19	
URTG DESRO	STANDARD; PRT; 394 AA.
ID	URTG DESRO STANDARD; PRT; 394 AA.
AC	P49150;
DT	01-FEB-1996 (Rel. 33, Created)
DT	01-FEB-1996 (Rel. 33, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Salivary plasminogen activator gamma precursor (EC 3.4.21.68) (DSPA gamma).
OS	Desmodus rotundus (Vampire bat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC	Desmodontinae; Desmodus.
OX	NCBI_TaxID=9430;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE-Salivary gland;
RX	MEDLINE=92039036; PubMed=1937019;
RA	Kraatzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA	Alagon A., Donner P., Schleuning W.D.;
RT	"The plasminogen activator family from the salivary gland of the vampire bat Desmodus rotundus: cloning and expression.";
RL	Gene 105:229-237(1991).
RN	[2]
RP	CHARACTERIZATION.
RX	MEDLINE=93393059; PubMed=1309059;
RA	Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA	Kraatzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA	Donner P.;
RT	"Plasminogen activators from the saliva of Desmodus rotundus (common vampire bat): unique fibrin specificity.";
RL	Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC	-!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC AGENT.
CC	-!- CATALYTIC ACTIVITY: Specific cleavage of Arg- -Val bond in plasminogen to form plasmin.

```

DR PIR; JS0600; JS0600.
DR HSSP; P98119; IASI.
DR MEROPS; S01.239; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; Signal; Multigene family.
FT SIGNAL 1 36 POTENTIAL.
FT CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT DOMAIN 45 126 KRINGLE.
FT DOMAIN 142 394 SERINE PROTEASE.
FT ACT_SITE 189 199 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 45 126 BY SIMILARITY.
FT DISULFID 66 108 BY SIMILARITY.
FT DISULFID 97 121 BY SIMILARITY.
FT DISULFID 131 262 BY SIMILARITY.
FT DISULFID 174 190 BY SIMILARITY.
FT DISULFID 182 251 BY SIMILARITY.
FT DISULFID 276 351 BY SIMILARITY.
FT DISULFID 308 324 BY SIMILARITY.
FT DISULFID 341 369 BY SIMILARITY.
FT CARBOHYD 315 315 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;

Query Match 28.1%; Score 223; DB 1; Length 394;
Best Local Similarity 44.3%; Pred. No. 1.2e-14;
Matches 39; Conservative 13; Mismatches 36; Indels 0; Gaps 0;

QY 45 DKSCTCYEGHFGYRKAGTDITMGRCPLPWSATVLQOQYHAHRSALQGLGKHNCRN 104
Db 40 DPHATCYKQGVYRTWTSTSGAGCINWNSLLIRRTYNGMPAVKVLGLGNHNYCRN 99

QY 105 PDNRPRWCYVQGLKPLVQECMVHDC 132
Db 100 PDGASKFCWCYVTKARKFTSESCSVPVCS 127

RESULT 20
FA12_BOVIN STANDARD; PRT; 593 AA.
AC P98140;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94342782; PubMed=816251;
RA Shibuya Y., Senba U., Okabe H., Kambara T., Yamamoto T.;
RT "Primary structure of bovine Hageman factor (blood coagulation factor

```

```

RT X11; comparison with human and guinea pig molecules.";
RL Biochim. Biophys. Acta 1206:63-70(1994).
RN [2]
RX SEQUENCE OF 10-21; 350-364 AND 525-550.
RX MEDLINE=77182112; PubMed=861210;
RA Fujikawa K., Walsh A.K., Davie W.E.;
RT "Isolation and characterization of bovine factor XII (Hageman
factor).";
RL Biochemistry 16:2270-2278(1977).
CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor Xla.
CC -!- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).
CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA. BOVINE FACTOR XII IS
CC CLEAVED ONLY TO ALPHA-FACTOR XIIA AS IT LACKS THE TRYPSIN/
CC KALLIKREIN CLEAVAGE SITE.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 2 EGF-like domains.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S70164; AAB30804.2; -.
DR PIR; S45281; S45281.
DR HSSP; P00763; LDPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrnctnl.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.

```


or send an email to license@isb-sib.ch).

```

CC  DR  EMBL; L20297; AAA28860.1; -.
CC  DR  EMBL; AJ003628; AAF52885.1; -.
CC  DR  EMBL; AJ002908; CRA05743.1; -.
CC  DR  PIR; A48289; A48289.
CC  DR  HSSP; P11362; IFGX.
CC  DR  FlyBase; Fgn0010407; Ror.
CC  DR  GO; GO:0016021; C:integral to membrane; NAS.
CC  DR  GO; GO:0004713; P:protein tyrosine kinase activity; NAS.
CC  DR  GO; GO:0007417; P:central nervous system development; IEP.
CC  DR  GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
CC  DR  InterPro; IPR000024; Fz domain.
CC  DR  InterPro; IPR000001; Kringle.
CC  DR  InterPro; IPR000719; Prot_kinase.
CC  DR  InterPro; IPR002011; RTKinaseII.
CC  DR  InterPro; IPR001245; Tyr_kinase.
CC  DR  Pfam; PF00051; kringle; 1.
CC  DR  Pfam; PF00069; pkinase; 1.
CC  DR  PRINTS; PR00018; KRINGLE.
CC  DR  PRINTS; PR00109; TYRKINASE.
CC  DR  ProDom; PD000395; Kringle; 1.
CC  DR  ProDom; PD000001; Prot_kinase; 1.
CC  DR  SMART; SM00130; KR; 1.
CC  DR  SMART; SM00219; Tyrc; 1.
CC  DR  PROSITE; PS50038; FZ; 1.
CC  DR  PROSITE; PS00021; KRINGLE_1; 1.
CC  DR  PROSITE; PS50070; KRINGLE_2; 1.
CC  DR  PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC  DR  PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC  DR  PROSITE; PS00109; PROTEIN KINASE TVR; 1.
CC  DR  PROSITE; PS00239; RECEPTOR TYR KIN II; 1.
CC  DR  Transferase; Kinase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW  Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW  Developmental protein.
FT  SIGNAL 1 24
FT  CHAIN 25 685
FT  DOMAIN 25 317
FT  TRANSMEM 318 338
FT  DOMAIN 339 685
FT  DOMAIN 36 225
FT  DOMAIN 236 310
FT  DOMAIN 410 677
FT  NP_BIND 416 424
FT  BINDING 442 442
FT  ACT_SITE 539 539
FT  MOD_RES 565 565
FT  MOD_RES 569 569
FT  MOD_RES 570 570
FT  CARBOHYD 45 45
FT  CARBOHYD 63 63
FT  CARBOHYD 129 129
FT  CARBOHYD 144 144
FT  CARBOHYD 250 250
FT  SEQUENCE 685 AA; 78142 MW; 526162D27D5FD7C7 CRC64;
SQ
Query Match 21.3%; Score 169; DB 1; Length 685;
Best Local Similarity 28.8%; Pred. No. 3.8e-09;
Matches 38; Conservative 27; Mismatches 35; Indels 32; Gaps 6;
QY 3 ELHQVPSNCDLNGGTCVSNKYFSNTHWCNPKFKGQHCHEIDKSKTCVEGNHFRGKA 62
DB 212 DOOKLPQHKDCSLGITI-----EVDKTCYWGDSGYRGVA 249
QY 63 STDTMGRCPLPMNSATVLOQTHAHRSADALQLGLGHKNYCRNPDN-RRRPWCYVQVGLKP 121
DB 250 NVSASGKPCLRW--SWLMKEI-----SDPEL-IGO-NYCRNPGSVENSPWCFVDSRR 300
QY 122 LVQECMVHDCAD 133

```

```

DB 301 ITELCDIPKCAD 312
RESULT 22
APOA_HUMAN
ID APOA_HUMAN STANDARD; PRT; 4548 AA.
AC P08519;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
CN LPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88039109; PubMed=3670400;
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA Fless G.M., Scanu A.M., Lawn R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RT plasminogen.";
RL Nature 330:132-137(1987).
RN [2]
RP SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibronectin and has serine proteinase
RT activity capable of cleaving it.";
RL EMBO J. 8:4035-4040(1989).
RN [3]
RP REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a).";
RL Science 246:904-910(1989).
RN [4]
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of o-glycans in conferring protease
RT resistance.";
RL J. Biol. Chem. 276:22200-22208(2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Iograsso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RT acid: existence of novel and expected binding modes.";
RL J. Mol. Biol. 256:751-761(1996).
RN [6]
RP VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scanu A.M., Pfaffinger D., Lee J.C., Hinnan J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNAcapha2-3Galbeta1-3GalNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.

```



```

RESULT 25
KRM1_XENLA
ID KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y90;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringie-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
[1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringie domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB070851; BAB4294.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringie.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringie; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PS00018; KRINGLE.
DR PRODOM; PD000395; Kringie; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00321; WSC; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR KX Wnt signaling pathway; Glycoprotein; Kringie;
KW Transmembrane; Signal.
FT SIGNAL 1 22
FT CHAIN 23 452
FT DOMAIN 23 369
FT TRANSMEM 370 390
FT DOMAIN 391 452
FT DOMAIN 29 112
FT DOMAIN 118 208
FT DOMAIN 212 319
FT CARBOHYD 43 43
FT CARBOHYD 57 57
FT CARBOHYD 215 215
FT CARBOHYD 253 253
FT CARBOHYD 291 291
FT CARBOHYD 328 328
FT CARBOHYD 344 344
FT CARBOHYD 452 452
SQ SEQUENCE 452 AA; 50198 MW; ED24BCD1AF4564E2 CRC64;

Query Match
Best Local Similarity 20.1%; Score 159; DB 1; Length 452;
Matches 31; Conservative 40.8%; Pred No. 2,4e-08;
42 CEIDKSKTCYGVNGHFYKRGKASTDTM--GRFCLPWSATVLIQOTYVHAHRSDALQLGLGKH 99

```

```

Db 22 CSDSPHSECYTVNGADYRGQTQNTSLDGGKPCLFWNE--TFQHPYNTLKYPNGEGLGKH 79
QY 100 NYCNPDPNRRPWCYV 115
Db 80 NYCNPDPDGVSPWCYI 95

RESULT 26
KRM1_MOUSE
ID KRM1_MOUSE STANDARD; PRT; 473 AA.
AC Q9NM43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringie-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.; DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
RA MEDLINE=21167372; PubMed=11267660;
RA Nakamura T.;
RL "Molecular cloning and characterization of Kremen, a novel
RL kringie-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
CC levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
CC on day 9 and increases up to day 18. Lower levels are found in
CC adult. At 9.5 dpc, expression is localised to the apical
CC ectodermal ridge (AER) of the developing fore- and hindlimb buds,
CC the telencephalon and the first brachial arch. At 10.5 dpc,
CC expression is also observed in the myotome and in sensory tissues
CC such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringie domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB059617; BAB40968.1; -.
DR HSPSP; P00747; 1CEA.
DR MGD; MGI:1933988; Kremen.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringie.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringie; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringie; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR KX Wnt signaling pathway; Signal; Transmembrane; Kringie.

```



```
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 315 315 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 549 Missing (in isoform Short).
FT SEQUENCE 937 AA; 104312 MW; 0D06945BF29F4773 CRC64;

Query Match 20.0%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 5.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGTGVSNKYFS-----NIHWCN-----PKKFGGQH 41
Db 236 DETSSVFKPRDLRCRDECEILENVLCQTEYIFARSNPMLRLKLPNCEDLPQESPEAAN 295

QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDAL 92
Db 296 CIRIGIPMADPINKNKHKNCYNSTGVDYRGVTSVTKSGRCQCPWNS-----QYPHTHSFTAL 350

QY 93 QLQ--LGKHNCRPNDRRR--PWCY 114
Db 351 RPEPLNGGHSYCRNPGNQKEAPWCF 375

RESULT 28
RORI_MOUSE
ID RORI_MOUSE STANDARD; PRT; 937 AA.
AC Q9Z139;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Tyrosine-protein kinase transmembrane receptor ROR1 precursor
DE (EC 2.7.1.112) (Neurotrophic tyrosine kinase, receptor-related 1)
DE (mRor1).
DE ROR1 OR NTRK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=92248426; PubMed=102311392;
RA Oishi I.; Takeuchi S.; Hashimoto R.; Nagabukuro A.; Ueda T.; Liu Z.J.;
RA Hattori.; Akira S.; Matsuda Y.; Yamamura H.; Otani H.; Minami Y.;
RT "Spatio-temporally regulated expression of receptor tyrosine kinases,
RT mRor1, mRor2, during mouse development: Implications in development
RT and function of the nervous system.";
RL Genes Cells 4:41-56(1999).
CC -!- FUNCTION: TYROSINE-PROTEIN KINASE RECEPTOR WHOSE ROLE IS NOT YET
CC CLEAR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. ROR
CC SUBFAMILY.
CC -!- SIMILARITY: Contains 1 frizzled (FZ) domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB010383; BAA75480.1; -.
CC HSSP; P00747; ICEA.
CC MGD; MGI:1347520; Ror1.
DR InterPro; IPR000024; Fz domain.
DR InterPro; IPR007110; Ig-like.
```

```
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000335; Kringle; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; Igc2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50038; FZ; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Receptor;
KW Transmembrane; Signal; Glycoprotein; Kringle; Phosphorylation;
KW Immunoglobulin domain.
FT SIGNAL 1 29
FT CHAIN 30 937
FT TYROSINE-PROTEIN KINASE TRANSMEMBRANE
FT RECEPTOR ROR1.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT IG-LIKE C2-TYPE.
FT FZ.
FT KRINGLE.
FT PROTEIN KINASE.
FT SER/THR-RICH.
FT PRO-RICH.
FT SER/THR-RICH.
FT NP_BIND 479 487
FT ATP (BY SIMILARITY).
FT BINDING 506 506
FT ATP (BY SIMILARITY).
FT ACT_SITE 615 615
FT MOD_RES 645 645
FT DISULFID 79 131
FT CARBOHYD 47 47
FT CARBOHYD 66 66
FT CARBOHYD 184 184
FT CARBOHYD 315 315
FT SEQUENCE 937 AA; 104156 MW; D728733E67D1782C CRC64;

Query Match 20.0%; Score 158.5; DB 1; Length 937;
Best Local Similarity 29.7%; Pred. No. 5.4e-08;
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7;

QY 2 NELHQP-----SNCDCLNGTGVSNKYFS-----NIHWCN-----PKKFGGQH 41
Db 236 DETSSVFKPRDLRCRDECEILENVLCQTEYIFARSNPMLRLKLPNCEDLPQESPEAAN 295

QY 42 C-----EIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHSDAL 92
Db 296 CIRIGIPMADPINKNKHKNCYNSTGVDYRGVTSVTKSGRCQCPWNS-----QYPHTHSFTAL 350

QY 93 QLQ--LGKHNCRPNDRRR--PWCY 114
Db 351 RPEPLNGGHSYCRNPGNQKEAPWCF 375

RESULT 29
KRM1_HUMAN
ID KRM1_HUMAN STANDARD; PRT; 475 AA.
AC Q96M08; Q9BY70; Q9UGS5; Q9UGU1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```


Search completed: December 3, 2003, 14:40:07
Job time : 5.33777 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 16.6105 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-4

Perfect score: 793

Sequence: 1 SNELHQVPSNCCLNGTGV.....QVGLKPLVQECWVHDCADGK 135

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

. Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	98.1	154	Q96SE8	Q96se8 homo sapien
2	602	75.9	433	Q8MIL0	Q8m10 oryctolagus
3	602	75.9	433	Q8MHI7	Q8mhy7 oryctolagus
4	550	69.4	157	Q9TV48	Q9tva8 bos taurus
5	548	69.1	231	Q8C6L2	Q8c6l2 mus musculus
6	322.5	40.7	516	Q9BU99	Q9bu99 homo sapien
7	315.5	39.8	559	Q91VP2	Q91vp2 mus musculus
8	313	39.5	214	Q9XT70	Q9xt70 oryctolagus
9	310.5	39.2	562	Q8SQ23	Q8sq23 sus scrofa
10	301.5	38.0	564	Q8MKB1	Q8mbk1 oryctolagus
11	264.5	33.4	517	Q8KOD2	Q8kod2 mus musculus
12	257.5	32.5	560	Q14520	Q14520 homo sapien
13	257	32.4	616	Q97507	Q97507 sus scrofa
14	250.5	31.6	653	Q8VCS4	Q8vcs4 mus musculus
15	249	31.4	597	Q35727	Q35727 mus musculus
16	233	29.4	615	Q81ZZ5	Q81zz5 homo sapien

17	229.5	28.9	395	4	Q9BZW1	Q9bzw1 homo sapien
18	218.5	27.6	90	4	Q8NGZ0	Q8ngz0 homo sapien
19	197.5	24.9	202	13	Q90675	Q90675 gallus gall
20	174	21.9	421	13	Q8AXX3	Q8axx3 xenopus lae
21	170.5	21.5	947	13	Q8AXY6	Q8axy6 gallus gall
22	166	20.9	716	13	Q91691	Q91691 xenopus lae
23	165.5	20.9	359	6	Q8WNR1	Q8wml1 canis famil
24	165	20.8	812	11	Q9R0W3	Q9r0w3 rattus norv
25	159	20.1	452	13	Q90Y90	Q90y90 xenopus lae
26	158.5	20.0	393	4	Q9BRB6	Q9brb6 homo sapien
27	158.5	20.0	937	11	Q8BNP9	Q8bnp9 mus musculu
28	158.5	20.0	937	11	Q8BG10	Q8bg10 mus musculu
29	158	19.9	454	6	Q46506	Q46506 papio hamad
30	156	19.7	113	4	Q9UIR5	Q9uirs homo sapien
31	154	19.4	806	6	Q18783	Q18783 macropus eu
32	154	19.4	810	4	Q15146	Q15146 homo sapien
33	153	19.3	113	4	Q9UIR7	Q9uir7 homo sapien
34	152.5	19.2	717	13	P70006	P70006 xenopus lae
35	151.5	19.1	648	4	Q9H1V4	Q9hlv4 homo sapien
36	150.5	19.0	567	4	Q13208	Q13208 homo sapien
37	150.5	19.0	930	13	Q8AV69	Q8av69 xenopus lae
38	149	18.8	801	11	Q8K0Q8	Q8k0q8 mus musculu
39	149	18.8	944	11	Q8C3W2	Q8c3w2 mus musculu
40	149	18.8	944	11	Q8BSP6	Q8bsp6 mus musculu
41	148.5	18.7	381	4	Q8N2J4	Q8n2j4 homo sapien
42	148.5	18.7	399	4	Q96GL8	Q96gl8 homo sapien
43	148.5	18.7	420	4	Q9BTP9	Q9btp9 homo sapien
44	148.5	18.7	424	4	Q8NCW1	Q8ncw1 homo sapien
45	148.5	18.7	462	4	Q8NCW0	Q8ncw0 homo sapien

ALIGNMENTS

RESULT 1

Q96SE8 PRELIMINARY; PRT; 154 AA.

AC Q96SE8
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Urokinase-type plasminogen activator amino-terminal fragment.
 GN ATF.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Fu J., Bai X., Ruan C.;
 RT "Cloning and expression of the amino-terminal fragment of human
 RT urokinase-type plasminogen activator,";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
 RT "Overexpression of the amino-terminal fragment of human urokinase-type
 RT plasminogen activator in breast cancer cells results in decreased
 RT tumor invasion, growth and angiogenesis,";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
 DR EMBL: AY029537; AAK38734.1; -.
 DR InterPro: IPR006209; EGF like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR000001; Kringle.
 DR Pfam: PF00051; kringle; 1.
 DR PRINTS: PR00018; KRINGLE.
 DR ProDom: PD000395; Kringle; 1.
 DR SMART: SM00181; EGF; 1.
 DR SMART: SM00130; KG; 1.
 DR PROSITE: PS00022; EGF 1; 1.
 DR PROSITE: PS00021; KRINGLE 1; 1.
 DR PROSITE: PS00070; KRINGLE_2; 1.


```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CC2FCFF505572 CRC64;

Query Match 98.1%; Score 778; DB 4; Length 154;
Best Local Similarity 99.3%; Pred. No. 1.8e-77;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFYRG 60
Db 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFYRG 80

Qy 61 KASDTMTGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQGLK 120
Db 81 KASDTMTGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQGLK 140

Qy 121 PLVQECMVHDCADG 134
Db 141 LLVQECMVHDCADG 154

RESULT 2
Q8MIL0 PRELIMINARY; PRT; 433 AA.
ID Q8MIL0
AC Q8MIL0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN PLAU.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22155945; PubMed=12149463;
RA Falkenberg M., Tom C., DeYoung M.B., Wen S., Linnemann R.,
RA Dichek D.A.;
RT "Increased expression of urokinase during atherosclerotic lesion
RT development causes arterial constriction and lumen loss, and
RT accelerates lesion growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10665-10670(2002).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR EMBL; AY122285; AAMB3187.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR PRINTS; PR0006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR PROSITE; PS00022; Tryp_Spc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48375 MW; 65B64F36415549B0 CRC64;

Query Match 75.9%; Score 602; DB 6; Length 433;
Best Local Similarity 76.6%; Pred. No. 1.1e-57;
Matches 105; Conservative 10; Mismatches 20; Indels 2; Gaps 1;

Qy 1 SNELHQV--PSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFY 58
Db 21 SNELHGVSDASNCGLNGGTCVTYKYFSNIWRCNPKFKFQGEHCEDTLTKCYHGDGHSY 80

Qy 59 RGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVG 118
Db 81 RGKANTDMDRPPCLAWNSANVLTQTYHAHRPDALQLGLGKHNYCRNPDHQRPPWCYVQVG 140

Qy 119 LKPLVQECMVHDCADGK 135
Db 141 LKQLIQECKVHDCSSGK 157

RESULT 3
Q8MHY7 PRELIMINARY; PRT; 433 AA.
ID Q8MHY7
AC Q8MHY7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX Sugiki M., Yoshida E., Anai K., Maruyama M.;
RN Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RT "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -
DR EMBL; AB087224; BAC02685.1; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match 75.9%; Score 602; DB 6; Length 433;
Best Local Similarity 76.6%; Pred. No. 1.1e-57;
Matches 105; Conservative 10; Mismatches 20; Indels 2; Gaps 1;

Qy 1 SNELHQV--PSNCDCLNGGTCVSNKYFSNIHWCNPKFGQHCEDKSKTCYEGNGHFY 58
Db 21 SNELHGVSDASNCGLNGGTCVTYKYFSNIWRCNPKFKFQGEHCEDTLTKCYHGDGHSY 80

Qy 59 RGKASTDTMGRPCLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPNRRRPPWCYVQVG 118
Db 81 RGKANTDMDRPPCLAWNSANVLTQTYHAHRPDALQLGLGKHNYCRNPDHQRPPWCYVQVG 140

Qy 119 LKPLVQECMVHDCADGK 135
Db 141 LKQLIQECKVHDCSSGK 157
```

[illegible]

QY 1 SNELH-----QVPSNCDLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTC 50
Db 23 SQEIHARFRGARSYGCCPRFCNGGTCCQALYFSDP-VCCPEGFAGKCCCEIDTRATC 81
QY 51 YEGNGHYFGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNRNR 110
Db 82 YEDQGISYRGTSWTAESGAECTWNSSALAKPYSGRRPDAIRLGLGNHNYCRNPDRD 141
QY 111 PWCYQVGLKPLVQECMVHDCADG 134
Db 142 PWCYVFKAGYSSEFCSTPACSEG 165

RESULT 7
Q91VP2 PRELIMINARY; PRT; 559 AA.
ID Q91VP2
AC Q91VP2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to plasminogen activator, tissue.
GN PLAT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; BC011256; AAH11256.1; -.
DR HSP; P00761; IAN1.
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibronectin.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. Protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp. SPC; 1.
DR PROSITE; PS00022; EGF 1; 1.
DR PROSITE; PS01185; EGF 2; 1.
DR PROSITE; PS01253; FIBRONECTIN 1; 1.
DR PROSITE; PS00021; KRINGLE 1; 2.
DR PROSITE; PS00070; KRINGLE 2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match 39.8%; Score 315.5; DB 11; Length 559;
Best Local Similarity 44.5%; Pred. No. 3.6e-26;
Matches 61; Conservative 15; Mismatches 56; Indels 5; Gaps 2;
QY 3 ELHQVP-----SNCDLNGGTCVSNKYFSNIHWCNPKKFGGQHCEIDKSKTCYEGNGHY 58
Db 74 QCHSVPRVSRSEPCRCFNGGTCCQALYFSDP-VCCPCPDGFGKRCIDTRATCFEEQGIT 132

QY 59 RGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNRNR 118
Db 133 RGTWTAESGAECINWSSVLKPYNARRPNAIKLGLGNHNYCRNPDRDLKPWCYVFK 192
QY 119 LKPLVQECMVHDCADGK 135
Db 193 GKYTEFCSTPACPKG 209

RESULT 8

Q9XT70 PRELIMINARY; PRT; 214 AA.
ID Q9XT70
AC Q9XT70
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Yin J.; Idell S.;
RL "Partial mRNA of rabbit uPA";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -|- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF097647; AAD39351.1; -.
DR HSP; P00749; IEGN.
DR MEROFS; S01.231; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. Protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp. SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
KW NON TER 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 39.5%; Score 313; DB 6; Length 214;
Best Local Similarity 78.3%; Pred. No. 2.4e-26;
Matches 54; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 67 MGRPCLPWNSATVLQOQTYHAHRSDALQGLGKHNCRPNRNR 126
Db 1 MDRPCLAWNSANVLTKTYHAHRPDALQGLGKHNCRPNRPHQR 60
QY 127 MVHDCADGK 135
Db 61 KVHDSGKK 69

RESULT 9

Q8SQ23 PRELIMINARY; PRT; 562 AA.
ID Q8SQ23
AC Q8SQ23
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSP; P00761; 1ANL.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00058; FN1; 1.
DR SMART; SM00136; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 562 AA; 63668 MW; F9B6B4C77CB101E8 CRC64;

Query Match 39.2%; Score 310.5; DB 6; Length 562;
Best Local Similarity 43.8%; Pred. No. 1.3e-25;
Matches 60; Conservative 14; Mismatches 58; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFY 58
Db 77 QCHSVPVKSCSPRCFNGTCTQAIYFSDF-VQCQPVGFGRQCEIDARATYEDGITY 135

QY 59 RGKASTDTMGRPCLPNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRWCYVQVG 118
Db 136 RGTWSTTESGAEVCNWNNTSGLASMPYNGRRPDAVKLGLGNHNYCRNPDKSKPWCYIFKA 195

QY 119 LKPLVQECMVHDCADGK 135
Db 196 EKYSDFCSTPACTREK 212

RESULT 10
Q8K0B1 ID Q8K0B1 PRELIMINARY; PRT; 564 AA.
AC Q8K0B1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AY029518; AAK40240.1; -.

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSP; P00761; 1ANL.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00058; FN1; 1.
DR SMART; SM00136; KR; 2.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 564 AA; 62726 MW; 459D8BAC6D4A937C CRC64;

Query Match 38.0%; Score 301.5; DB 6; Length 564;
Best Local Similarity 48.7%; Pred. No. 1.2e-24;
Matches 56; Conservative 13; Mismatches 41; Indels 5; Gaps 3;

QY 5 HQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCIEDKSKTCYEGNGHFYRG 60
Db 80 HSPVQSCSEPRCLNGGTCQAIYFSDF-VQCQPEGFVGRCEVDTRARCYEDRGIGYRG 138

QY 61 KASTTMTGRPCLPNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDNRPRWCYV 115
Db 139 TWSSTTESGACQVNNWSNLALPKPYSGRKPNALRLGLGNHNYCRNPDRTKPCYV 193

RESULT 11
Q8K0D2 ID Q8K0D2 PRELIMINARY; PRT; 517 AA.
AC Q8K0D2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
```



```
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase;
KW Serine protease.
SQ SEQUENCE 616 AA; 68012 MW; 4C5FE3D71EBBDIA9 CRC64;

Query Match 32.4%; Score 257; DB 6; Length 616;
Best Local Similarity 40.9%; Pred. No. 1e-19;
Matches 52; Conservative 14; Mismatches 53; Indels 8; Gaps 3;

QY 6 QVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFGYRGKASTD 65
DB 176 QVCSINPLNGSGCLQTE---GHRLCRCPGYAGRLCDVDLKERCYSDRGLSYRGMAQIT 232

QY 66 TWGRPCLPNNSATVLQOITY-HAHRSDALQILGLGKHNYCRNPDRRPPWCYVQVGLKPLVQ 124
DB 233 LSGAPQCPWAS----EATYWNMTAEQALNWLGDHAFRCRPNPDTRPCWFCVWRGDQLSWQ 288

QY 125 ECMVHDC 131
DB 289 YKRLARC 295

RESULT 14
Q8VCS4 PRELIMINARY; PRT; 653 AA.
AC Q8VCS4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2003 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; 1AN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR000722; CHYMOTRYPsin.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD00018; KRINGLE.

Query Match 31.6%; Score 250.5; DB 11; Length 653;
Best Local Similarity 42.0%; Pred. No. 5.6e-19;
Matches 50; Conservative 11; Mismatches 41; Indels 17; Gaps 2;

QY 5 HQVPSNCDCLNGGTCVSNKYFSNIHW-----CNCPPKFGGQHCIDKSKTCYEGNGH 56
DB 239 HTACLSPCLNGGTC-----HLIVGTGTSVCTCPLGYAGRFPNIVPTEHFLNGT 289

QY 57 FYRGKASTDTMGRCPLPWNNSATVLQOITYHAHRSDALQILGLGKHNYCRNPDRRPPWCYV 115
DB 290 EYRGVASTAASGLSCLAMNSDLLYQELHVDVSAVAAILGLGHAYCRNPDKDRPPWCYV 348

RESULT 15
Q35727 PRELIMINARY; PRT; 597 AA.
AC Q35727;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Schloesser M.; Schwager S.; Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X99571; CAA67891.1; -.
DR HSSP; P00760; 1AQ7.
DR MEROPS; S01.211; -.
DR MGD; MGI.1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR000722; CHYMOTRYPsin.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD00018; KRINGLE.
DR ProDom; PD000995; FN Type II; 1.
DR ProDom; PD000395; Kringle; 1.
```

```

DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FBA CRC64;

Query Match 31.4%; Score 249; DB 11; Length 597;
Best Local Similarity 40.8%; Pred. No. 7.4e-19;
Matches 49; Conservative 14; Mismatches 49; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKPSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
183 CLNGGSC-----LVEDHPLCRCPYGYGFCDLDMATCYEGRGLSYRGAGTTQSGAPCQ 239

QY 73 PMNSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDNRPRWCYVQVGLKPLVOECMVHDC 131
240 RW-----TVEATYRNMTKEQALSGWGLGHAFCRPNPDNDTRPWCFCVWSGDRLSWDYCGLEQC 295

RESULT 16
Q81Z25 PRELIMINARY; PRT; 615 AA.
AC Q81Z25;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SQ SEQUENCE FROM N.A.

Query Match 29.4%; Score 233; DB 4; Length 615;
Best Local Similarity 42.3%; Pred. No. 4.3e-17;
Matches 44; Conservative 12; Mismatches 40; Indels 8; Gaps 3;

QY 13 CLNGGTCVSNKPSNIHWCNCPKFGQHCHEIDKSKTCYEGNGHFYRGKASTDTMGRPCL 72
183 CLNGGRCL---VEGHLRCLCPVGYTFDVTDKASKYDGRGLSYRGLARTILSGAPCQ 239

QY 73 PMNSATVLOQTY-HAHRSDALQLGLGKHNYCRNPDNRPRWCYV 115
240 PWAS-----EATYRNVTAEQARNWGLGGHAFCRPNPDNDIRPWCYV 279

RESULT 17
Q9B2W1 PRELIMINARY; PRT; 395 AA.
AC Q9B2W1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).
OS Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; 1PK2
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02440; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 28.9%; Score 229.5; DB 4; Length 395;
Best Local Similarity 42.7%; Pred. No. 6.5e-17;
Matches 44; Conservative 10; Mismatches 44; Indels 5; Gaps 2;

QY 30 WCNCPKFGQHCHEIDKSKTCYEGNGHFYRGKASTDTMGRPCLPWSATVLOQTYHAHRS 89
33 WCNCS---GRAQCS-EGNSDCYFGNGSAVRGTHSLTESGASCLPWSMILIGKVYTAQNP 87

QY 90 DALQLGLGKHNYCRNPDNRPRWCYVQVGLKPLVOECMVHDC 132
88 SAQALGLGKHNYCRNPDGDAKPMWCHLVKNRRLTWECYCDVPSCS 130

RESULT 18
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "Production of kringle fragment.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF282882; AAM52248.1; -.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.

```

DR PROS:TE; P850070; KRINGLE_2; 1.
 SW Glycoprotein; Kringle.
 SQ SEQUENCE 90 AA; 9804 MW; A33887F9DF4C7B1 CRC64;
 Query Match 27.6%; Score 218.5; DB 4; Length 90;
 Best Local Similarity 49.4%; Pred. No. 2.1e-16;
 Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;
 QY 50 CYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPDR 109
 Db 8 CYFGNGSAYRGTHSLTESGASCLPWNMILIGKVITQAQPSAQAQGLGKHNYCRNP 67
 QY 110 RWCYVQVGLKPLVQECMHDC 132
 Db 68 KPCWYT-TNPRKLYDYCDVPOCA 89
 RESULT 19
 Q90675 PRELIMINARY; PRT; 202 AA.
 ID Q90675; PRELIMINARY; PRT; 202 AA.
 AC Q90675;
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Tissue-type plasminogen activator (Fragment).
 GN TPA.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=White Leghorn;
 RC MEDLINE=97199025; PubMed=9047000;
 RA Johnson A.L., Brigham J.T., Anthony R.V.;
 RT "Expression of avian urokinase and tissue-type plasminogen activator
 RT messenger ribonucleic acid during follicle development and atresia.";
 RL Biol. Reprod. 56:581-588(1997).
 CC -1- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
 DR EMBL; U31988; AAA74955.1; --
 DR HSSP; P00750; 1RTF.
 DR MEROPS; S01.234; --
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00018; KRINGLE
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00130; KG; 1.
 DR PROSITE; PS00021; KRINGLE 1; 1.
 DR PROSITE; PS00070; KRINGLE 2; 2.
 DR PROSITE; PS02400; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 202
 SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;
 Query Match 24.9%; Score 197.5; DB 13; Length 202;
 Best Local Similarity 36.9%; Pred. No. 1e-13;
 Matches 41; Conservative 11; Mismatches 50; Indels 9; Gaps 2;
 QY 30 WNCNPK--KFGQHC-----IDSKTKCYEGNGHFGYRGKASTDTMGRPCLPWNSATVL 80
 Db 12 WCYVFAGKVISEFCSTPACTKVAEDGDCYTGNGIAGVGRSTRYSKSGFCLPWNVPFLT 71
 QY 81 QOYTHAHRSDALQGLGKHNYCRNPDRRRPCWYQVGLKPLVQECMHDC 131
 Db 72 SKIYTALEQRALGKGNHCRNPDGDAQPCWCHWKDRQLTWECYDVPQC 122
 RESULT 20
 Q90675 PRELIMINARY; PRT; 421 AA.
 ID Q90675; PRELIMINARY; PRT; 421 AA.
 AC Q90675;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Kremen2.
 GN KRM2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22308873; PubMed=12421700;
 RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
 RT "Kremen proteins interact with Dickkopf1 to regulate anteroposterior
 RT CNS patterning";
 RL Development 129:5587-5596(2002).
 DR EMBL; AY150813; AAN64661.1; --
 SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;
 Query Match 21.9%; Score 174; DB 13; Length 421;
 Best Local Similarity 47.4%; Pred. No. 8.4e-11;
 Matches 36; Conservative 6; Mismatches 28; Indels 6; Gaps 3;
 QY 41 HCEIDKSKTCYEGNGHFGYRGKAS-TDTMGRPCLPWNSATVLQOYTHAHRSDALQGLGKH 99
 Db 23 HPFLSE---CFVNGRDYRGVTSQAGPEGTCLYWNQTT--QHLYNQSDPDGELGNH 77
 QY 100 NYCRRNPDRRRPCWYV 115
 Db 78 NYCRRNPDAVDQPCWYV 93
 RESULT 21
 Q90675 PRELIMINARY; PRT; 947 AA.
 ID Q90675; PRELIMINARY; PRT; 947 AA.
 AC Q90675;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Muscle-specific receptor tyrosine kinase MusK.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=20538710; PubMed=11083926;
 RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.K.,
 RA Yancopoulos G.D., Ip N.Y.;
 RT "Cloning and characterization of muscle-specific kinase in chicken";
 RL Mol. Cell. Neurosci. 16:661-673(2000).
 CC [2]
 RN SEQUENCE FROM N.A.
 RP Gies D., Glass D.J., Yancopoulos G.D.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY143173; AAN05008.1; --
 KW Receptor; Kinase.
 SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;
 Query Match 21.5%; Score 170.5; DB 13; Length 947;
 Best Local Similarity 32.4%; Pred. No. 4.9e-10;
 Matches 46; Conservative 13; Mismatches 34; Indels 49; Gaps 8;
 QY 4 LHOVSNCDLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFGYRGKAS 63
 Db 441 LHQDPSAC-----THIFFDFKKE-----NITRTCYSGNGGQFYQGAN 478
 QY 64 TDTMGRPCLPWNSATVLQOYTHAHR-----SDALQGLGKHNYCRNP-DNRRPWC 113


```
Db 479 VTASGIPCKWS-----DOAPHLHRRTPQVPPELSDA-----ENYCRNPGGENRPMC 526
Qy 114 YVQVGLKPLV--QECMVHDCAD 133
Db 527 YTK---DPSWTWYCVSVPDGD 545

RESULT 22
Q91691 PRELIMINARY; PRT; 716 AA.
AC Q91691;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Growth factor Livertine.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ruiz i Altaba A., Thery C.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR HSSP; P00747; 1CEA.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan_app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 4.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 716 AA; 81971 MW; 508376A0E4398798 CRC64;

Query Match 20.9%; Score 166; DB 13; Length 716;
Best Local Similarity 31.1%; Pred. No. 1.1e-09;
Matches 42; Conservative 11; Mismatches 40; Indels 42; Gaps 6;

Qy 30 WC-----NCPKFGQGHCEIDKSK-----TCYEGNGHFYRGKASTDWMGRCLP 73
Db 258 WCYTDDNVEKEP-----CRITCKKQRLSNIETSTCFKRGEGYRGKANTTSGIPQOR 313
Qy 74 WNSAT-----VLQOTYHAHRSALQLGLGKHNYCRNPDNRNRRPWCYVQV-----GL 119
Db 314 WDSQTPQSHRFLPEKYPCKGLD-----ENYCRNPDGSEAPWCFTTLPGRMAYCFQI 365
Qy 120 KPLVQECMVHDCADG 134
Db 366 KRCKDDVLEPDCVHG 380

RESULT 23
Q8WNR1 PRELIMINARY; PRT; 359 AA.
AC Q8WNR1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
```

```
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069985; AAL58519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 4.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS00070; KRINGLE_2; 4.
KW Glycoprotein; Kringle.
FT NON_TER 359
FT SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;

Query Match 20.9%; Score 165.5; DB 6; Length 359;
Best Local Similarity 29.0%; Pred. No. 6.1e-10;
Matches 42; Conservative 13; Mismatches 43; Indels 47; Gaps 7;

Qy 5 HQVPSNCDCLNGTGVSNKYFSN-----IHWCNCPKFGQGHCEIDKSKT-----49
Db 211 NRTFENFPCKN-----LDENYCRNPDGETAPWCYTTSNEVEHQCIPSCSSPITTEYLD 266
Qy 50 -----CYEGNGHFYRGKASTDWMGRCLPWN SATVLQOTYHAHRSAL--92
Db 267 APASVPPEQTPVVQECYHGNGQSYRGTSSTITGRKQCSWSMT-----PHRHEKTP 319
Qy 93 ---QLGLGKHNYCRNPDNRNRRPWCY 114
Db 320 HPFEAGL-TWNYCRNPDADKSPWCY 343

RESULT 24
Q9ROW3 PRELIMINARY; PRT; 812 AA.
AC Q9ROW3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
GN PLASMINOGEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Bangert K., Johnsen A.H., Thorsen S.;
RT "Rat plasminogen: cDNA and gene structure.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91250378; PubMed=1645711;
RA Kanakas J.J.; Makker S.P.;
RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
RT receptor site for plasminogen.";
RL J. Biol. Chem. 266:10825-10829(1991).
```


Db	236	DETSSVPKPRDLCRDCEILNVLQCTYIIFARSNPMLRLKLPNCEDLPQESPEAAN	295
Qy	42	C-----EIDSKTCYEGNGHFYRGKASTDTMGRECLPWNSATVLQOTYHAHRSDAL	92
Db	296	CIRIGIPWADPINKNHKCYNSTGVDYRGTVSVTKSGRCQCPWNS-----QYPHTHTFTAL	350
Qy	93	QLG--LGKHNYCRNPDRRR--PCWCY 114	
Db	351	RFPELNGGHSYCRNPGNKEAPWCF 375	
RESULT 27			
Q8BNP9		PRELIMINARY;	PRT; 937 AA.
ID	Q8BNP9		
AC	Q8BNP9		
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	Receptor tyrosine kinase-like orphan receptor 1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
[1]			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Adipose tissue;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RL	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
DR	EMBL; AK080896; BAC38068.1; --		
SQ	SEQUENCE 937 AA; 104118 MW; CB440323CA66EC28 CRC64;		
Query Match 20.0%; Score 158.5; DB 11; Length 937;			
Best Local Similarity 29.7%; Pred. No. 1e-08;			
Matches 43; Conservative 14; Mismatches 51; Indels 37; Gaps 7			
Qy	2	NELHQVP-----SNCDLNGGTGVSNKYFS-----NIHWNCN-----PKKFGGQH 41	
Db	236	DETSSVPKPRDLCRDCEVLENVLCQTYIFARSNPMLRLKLPNCEDLPQESPEAAN	295
Qy	42	C-----EIDSKTCYEGNGHFYRGKASTDTMGRECLPWNSATVLQOTYHAHRSDAL	92
Db	296	CIRIGIPWADPINKNHKCYNSTGVDYRGTVSVTKSGRCQCPWNS-----QYPHTHTFTAL	350
Qy	93	QLG--LGKHNYCRNPDRRR--PCWCY 114	
Db	351	RFPELNGGHSYCRNPGNKEAPWCF 375	
RESULT 28			
Q8BG10		PRELIMINARY;	PRT; 937 AA.
ID	Q8BG10		
AC	Q8BG10		
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-MAR-2003	(TREMBLrel. 23, Last annotation update)	
DE	Receptor tyrosine kinase-like orphan receptor 1.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
[1]			
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Adipose tissue;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		

Db 102 CYHGDCQSYRGSGFSTVTGRTCSWSSMTFPHQHKRTPENHPNDGLTM-----NYCRNPD 156

QY 108 RRRPWCYVQVGLKPLV--QECMVHDCAD 133

Db 157 DTGPGWCF--MDPSVRWEYCNLTRCSD 181

RESULT 30

Q9UIRS PRELIMINARY; PRT; 113 AA.

AC Q9UIRS;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Apolipoprotein(a) (Fragment).

GN APOA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21181705; PubMed=11285247;

RA Ogorekova M., Kraft H.G., Enholm C., Utermann G.;

RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV

RT types 6 to 10 domain affect lip(a) plasma concentrations and have

RT different patterns in Africans and Caucasians.";

RL Hum. Mol. Genet. 10:815-824(2001).

CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF158663; AAF03680.1; -.

DR EMBL; AF158662; AAF03680.1; JOINED.

DR HSSP; P00747; IPMK.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; kringle; 1.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00021; KRINGLE_1; 1.

DR PROSITE; PS00070; KRINGLE_2; 1.

KW Glycoprotein; Kringle; Lipoprotein.

FT NON_TER 1 1

FT NON_TER 113 113

SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match 19.7%; Score 156; DB 4; Length 113;

Best Local Similarity 37.8%; Pred. No. 1.9e-09;

Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps 4;

QY 48 KTCYEGNGHFYRGKASDTMGRPCLPWNSATV--LQCTYHAHESDALQLGLGKHNYCRNP 105

Db 9 RCYHNGNGSYRGTFSTVTGRTCSWSSMTFPHQHKRTPENHPNDGLTM-----NYCRNP 63

QY 106 DNRRRWCYVQVGLKPLV--QECMVHDCAD 133

Db 64 DADTGPWCFT--MDPSIRWEYCNLTRCSD 90

Search completed: December 3, 2003, 14:42:57

Job time : 16.6105 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 13.8597 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-4
Perfect score: 793
Sequence: 1 SNEHQVPSNCDCLNGGTCV.....QVGLKPLVQCMVHDCADGK 135

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793	100.0	135	10	US-09-880-503-4
2	793	100.0	138	10	US-09-884-186-12
3	793	100.0	138	15	US-10-237-667-12
4	793	100.0	138	15	US-10-237-708-12
5	793	100.0	138	15	US-10-237-866-12
6	793	100.0	138	15	US-10-237-871-12
7	793	100.0	138	15	US-10-237-624-12
8	793	100.0	143	10	US-09-880-503-8
9	793	100.0	337	15	US-10-106-698-6266
10	793	100.0	403	10	US-09-880-503-6
11	793	100.0	411	10	US-09-880-503-3
12	793	100.0	431	10	US-09-264-468B-1
13	793	100.0	431	12	US-10-301-822-161
14	793	100.0	431	12	US-10-131-985-21
15	793	100.0	431	14	US-10-076-421-2

16	793	100.0	431	15	US-10-171-311-184	Sequence 184, Appl
17	793	100.0	431	15	US-10-193-656-4	Sequence 4, Appl
18	791	99.7	411	12	US-10-247-671-149	Sequence 149, Appl
19	783	98.7	411	12	US-10-407-821-2	Sequence 2, Appl
20	510	64.3	88	10	US-09-880-503-1	Sequence 1, Appl
21	510	64.3	96	10	US-09-880-503-9	Sequence 9, Appl
22	328.5	41.4	527	11	US-09-987-457-18	Sequence 18, Appl
23	328.5	41.4	527	11	US-09-987-455-19	Sequence 19, Appl
24	328.5	41.4	562	9	US-09-969-271-7	Sequence 7, Appl
25	328.5	41.4	562	10	US-09-974-298-145	Sequence 145, Appl
26	328.5	41.4	562	12	US-10-443-701-4	Sequence 4, Appl
27	328.5	41.4	562	15	US-10-193-656-8	Sequence 8, Appl
28	293	36.9	49	12	US-10-349-543-5	Sequence 5, Appl
29	288	36.3	323	10	US-09-880-503-7	Sequence 7, Appl
30	283	35.7	47	10	US-09-880-503-2	Sequence 2, Appl
31	268	33.8	44	12	US-10-349-543-1	Sequence 1, Appl
32	260.5	32.8	655	15	US-10-172-712-28	Sequence 28, Appl
33	257.5	32.5	560	10	US-09-912-559-3	Sequence 3, Appl
34	257.5	32.5	560	10	US-09-912-559-4	Sequence 4, Appl
35	257.5	32.5	560	15	US-10-172-712-32	Sequence 32, Appl
36	233	29.4	615	11	US-09-858-909-2	Sequence 2, Appl
37	233	29.4	615	15	US-10-172-712-30	Sequence 30, Appl
38	220.5	27.8	326	14	US-10-057-951-3	Sequence 3, Appl
39	220.5	27.8	372	9	US-09-084-491A-3	Sequence 3, Appl
40	220.5	27.8	372	14	US-10-102-704-3	Sequence 3, Appl
41	215	27.1	354	11	US-09-987-457-10	Sequence 10, Appl
42	215	27.1	354	11	US-09-987-455-11	Sequence 11, Appl
43	215	27.1	377	11	US-09-987-455-8	Sequence 8, Appl
44	176	22.2	343	11	US-09-987-457-14	Sequence 14, Appl
45	176	22.2	343	11	US-09-987-455-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-4

Query Match		100.0%	Score 793;	DB 10;	Length 135;
Best Local Similarity		100.0%	Pred. No. 2.6e-72;		
Matches 135;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG	60		
Db	1	SNEHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG	60		
Qy	61	KASDTMTGRCLPWSATVLQOTYHAHRSALQGLGKHNYCRPNDRRRPWCVVQVGLK	120		
Db	61	KASDTMTGRCLPWSATVLQOTYHAHRSALQGLGKHNYCRPNDRRRPWCVVQVGLK	120		
Qy	121	PLVQECMVHDCADGK	135		
Db	121	PLVQECMVHDCADGK	135		

RESULT 2

```

US-09-984-186-12
; Sequence 12, Application US/09984186
; Patent No. US2002015101A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL
; CONTAINING SAID POLYPEPTIDES
;

```

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (Patentin)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/984,186
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>

CLASSIFICATION: CONFIDENTIAL
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCY/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: SY92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

RESULT 3

```

US-10-237-667-12
; Sequence 12, Application US/10237657
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittou, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES, COMPOSITION
; PREPARATION THEREOF AND PHARMACEUTICAL
; CONTAINING SAID POLYPEPTIDES
;

```

NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: System 7.1
SOFTWARE: Word 5.1 (PatentIn)
CURRENT APPLICATION DATA:
APPLICANT APPLICATION NUMBER: US/10/237,667
FILING DATE: 10-Sep-2002
CLASSIFICATION: <Unknown>

PRIOR CLASSIFICATION DATA:
APPLICATION NUMBER: US/08/797,689
FILING DATE: 31-JAN-1997
APPLICATION NUMBER: US 08/256,927
FILING DATE: 28-JUL-1994
APPLICATION NUMBER: FR 92/01064
FILING DATE: 31-JAN-1992
APPLICATION NUMBER: PCT/FR93/00085
FILING DATE: 28-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Smith Ph.D., Julie K.
REGISTRATION NUMBER: P-38,619
REFERENCE/DOCKET NUMBER: ST92006-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3839
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-667-12

RESULT 4

```
US-10-237-708-12
; Sequence 12, Application US/10237708
; Publication No. US20030036170A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
;             Fournier, Alain
;             Guitton, Jean-Dominique
;             Jung, Gerard
;             Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,708
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-708-12
Query Match 100.0%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.6e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
Qy 61 KASDTMTGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 64 KASDTMTGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138
RESULT 5
US-10-237-866-12
; Sequence 12, Application US/10237866
; Publication No. US20030036171A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
;             Fournier, Alain
;             Guitton, Jean-Dominique
;             Jung, Gerard
;             Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,866
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12
Query Match 100.0%; Score 793; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.6e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 60
Db 4 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGHFYRG 63
Qy 61 KASDTMTGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 120
Db 64 KASDTMTGRCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK 123
Qy 121 PLVQECMVHDCADGK 135
Db 124 PLVQECMVHDCADGK 138
RESULT 6
US-10-237-871-12
; Sequence 12, Application US/10237871
```

Publication No. US20030036172A1

GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 Fournier, Alain
 Guittou, Jean-Dominique
 Jung, Gerard
 Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/237,871
 FILING DATE: 10-Sep-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/797,689
 FILING DATE: 31-JAN-1997
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-JUL-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-JAN-1992
 APPLICATION NUMBER: PCT/FR93/00085
 FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Smith Ph.D., Julie K.
 REGISTRATION NUMBER: P-38,619
 REFERENCE/DOCKET NUMBER: ST92006-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 138 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-237-871-12

Query Match 100.0%; Score 793; DB 15; Length 138;
 Best Local Similarity 100.0%; Pred. No. 2.6e-72;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDLNGGTCVSNKYFSNTHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
 Db 4 SNELHVPNSCDLNGGTCVSNKYFSNTHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 63

QY 61 KASTDTMGRECLPWNASATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRRCWCVQVGLK 120
 Db 64 KASTDTMGRECLPWNASATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRRCWCVQVGLK 123

QY 121 PLVQECMVHDCADGK 135
 Db 124 PLVQECMVHDCADGK 138

RESULT 7
 US-10-237-624-12
 ; Sequence 12, Application US/10237624
 ; Publication No. US20030082747A1

Publication No. US20030036172A1

GENERAL INFORMATION:
 APPLICANT: Fleer, Reinhard
 Fournier, Alain
 Guittou, Jean-Dominique
 Jung, Gerard
 Yeh, Patrice

TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
 PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
 CONTAINING SAID POLYPEPTIDES

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3C43
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Macintosh
 OPERATING SYSTEM: System 7.1
 SOFTWARE: Word 5.1 (PatentIn)

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/237,871
 FILING DATE: 10-Sep-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/797,689
 FILING DATE: 31-JAN-1997
 APPLICATION NUMBER: US 08/256,927
 FILING DATE: 28-JUL-1994
 APPLICATION NUMBER: FR 92/01064
 FILING DATE: 31-JAN-1992
 APPLICATION NUMBER: PCT/FR93/00085
 FILING DATE: 28-JAN-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Smith Ph.D., Julie K.
 REGISTRATION NUMBER: P-38,619
 REFERENCE/DOCKET NUMBER: ST92006-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610) 454-3839
 TELEFAX: (610) 454-3808

INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 138 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 12:
 US-10-237-871-12

Query Match 100.0%; Score 793; DB 15; Length 138;
 Best Local Similarity 100.0%; Pred. No. 2.6e-72;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHVPNSCDLNGGTCVSNKYFSNTHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 60
 Db 4 SNELHVPNSCDLNGGTCVSNKYFSNTHWCNCPKFGGQHCEIDKSKTCYEGNGHYRG 63

QY 61 KASTDTMGRECLPWNASATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRRCWCVQVGLK 120
 Db 64 KASTDTMGRECLPWNASATVLOQTYHAHRSALQLGLGKHNYCRNPNRRRRCWCVQVGLK 123

QY 121 PLVQECMVHDCADGK 135
 Db 124 PLVQECMVHDCADGK 138

RESULT 7
 US-10-237-624-12
 ; Sequence 12, Application US/10237624
 ; Publication No. US20030082747A1

; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match 100.0%; Score 793; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 2.7e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 9

US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match 100.0%; Score 793; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 6.9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 27 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 86
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 87 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 146
QY 121 PLVQECMVHDCADGK 135
DB 147 PLVQECMVHDCADGK 161

RESULT 10

US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 100.0%; Score 793; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 8.3e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
DB 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120
QY 121 PLVQECMVHDCADGK 135
DB 121 PLVQECMVHDCADGK 135

RESULT 11

US-09-880-503-3
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; FILE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match 100.0%; Score 793; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 8.5e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
DB 1 SNELHQVPSNCDCLNGTGVSNKYFSNIHWCNCPKFGGQHCIDKSKTCYEGNGHFYRG 60
QY 61 KASTDTMGRPCLPWNSATVLQOITYHAHRSDALQGLGKHNYCRPNDRRPPWCYVQVGLK 120

Db 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120

Qy 121 PLVQECMVHDCADGK 135
Db 121 PLVQECMVHDCADGK 135

RESULT 12

US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310.US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match 100.0%; Score 793; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80

Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155

RESULT 13

US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.

; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match 100.0%; Score 793; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80

Qy 61 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 120
Db 81 KASTDTMGRPCLPWN SATVLQOTYHAHRSDALQLGLGKHNYCRNPDNRPRPCYVQVGLK 140

Qy 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155

RESULT 14

US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occlleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match 100.0%; Score 793; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 9e-72;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 60
Db 21 SNELHQPNSCDCLNGGTCVSNKYFSNIHWCNCPKFGQHCIDKSKTCYEGNGHFYRG 80

Qy	61	KAS	TD	TW	GR	PC	LW	NS	AT	VL	QO	TH	AH	RS	DA	LQ	LG	KN	YCR	PN	DR	RR	PW	CY	VO	VL	120
Dd	81	KAS	TD	TW	GR	PC	LW	NS	AT	VL	QO	TH	AH	RS	DA	LQ	LG	KN	YCR	PN	DR	RR	PW	CY	VO	VL	140

Qy 121 PLVQECMVHDCADGK 135
Db 141 PLVQECMVHDCADGK 155

```

RESULT 15
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAOKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

```

Query Match	100.0%	Score 793;	DB 14;	Length 431;
Best Local Similarity	100.0%	Pred. No. 9e-72;		
Matches 135;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
D6	S	N	E	L	H	Q	V	P	N	C	D	C	L	I	N	G	T	C	V	S	N	K	Y	F	S	N	I	H	W	C	N	C	P	K	F	G	G	O	H	C	E	I	D	S	K	T	C	V	E	G	N	G	H	F	Y	R	G			
D6	S	N	E	L	H	Q	V	P	N	C	D	C	L	I	N	G	T	C	V	S	N	K	Y	F	S	N	I	H	W	C	N	C	P	K	F	G	G	O	H	C	E	I	D	S	K	T	C	V	E	G	N	G	H	F	Y	R	G			

Qy	61	KASDTDMGPPCLPWN SATVLQOTYHAHRS DALOLGLGHKNYCRNPDRRRRWCVVQVGLK	120
Db	81	KASDTDMGRPCLPWN SATVLQOTYHAHRS DALQLGLGHKNYCRNPDRRRRWCVVQVGLK	140

Qy 121 PLVQECMVHDCADGK 135
|||
Db 141 PLVQECMVHDCADGK 155

```

RESULT 16
US/10-171-311-184
> Sequence 184, Application US/10171311
> Publication No. US2003008270A1
> GENERAL INFORMATION:
> APPLICANT: Schlegel, Robert
> APPLICANT: Chen, Yan
> APPLICANT: Zhao, Xumei
> APPLICANT: Monahan, John
> APPLICANT: Kamatkar, Shubhangi
> APPLICANT: Glatt, Karen
> APPLICANT: Gannavarapu, Manjula
> APPLICANT: Hoersch, Sebastian
> TITLE OF INVENTION: NOVEL GENES, COM
> TITLE OF INVENTION: IDENTIFICATION,
> TITLE OF INVENTION: OF CERVICAL CAN
> FILE REFERENCE: MRI-035
> CURRENT APPLICATION NUMBER: US/10/17
> CURRENT FILING DATE: 2002-06-12
> PRIOR APPLICATION NUMBER: US 60/298,
> PRIOR FILING DATE: 2001-06-13
> PRIOR APPLICATION NUMBER: US 60/298,
> PRIOR FILING DATE: 2001-06-13

```

```

; PRIOR APPLICATION NUMBER: US 60/335,936
;
; PRIOR FILING DATE: 2001-11-14
;
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

```

Query Match	100.0%;	Score 793;	DB 15;	Length 431;
Best local Similarity	100.0%;	Pred. No. 9e-72;		
Matches 135;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGH	FVRG	60
Db	21	SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCEIDKSKTCYEGNGH	FVRG	80
Qy	61	KASDTDTWGRFCLPWSATVLQOTYHAHRSDALQLGLGKHNRCRPNDRRRPWCYVQVGLK	120	
Db	81	KASDTDTWGRFCLPWSATVLQOTYHAHRSDALQLGLGKHNRCRPNDRRRPWCYVQVGLK	140	
Qy	121	PLVOECMVHDCADGK	135	
Db	141	PLVOECMVHDCADGK	155	

```

RESULT 17
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733a1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: Li, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

```

	Query Match	100.0%	Score 793;	DB 15;	Length 431;
	Best Local Similarity	100.0%	Pred. No. 9e-72;		
	Matches 135;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	SNELHQVPSNCDCLNGATCTVSNKYFSTHWCNCPKFKGGQHCETDKSKTCYEGNGHFFYRG	60		
Db	21	SNELHQVPSNCDCLNGATCTVSNKYFSTHWCNCPKFKGGQHCETDKSKTCYEGNGHFFYRG	80		
Qy	61	KASDTDTMGRFCLPNSATVLIQOTVHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK	120		
Db	81	KASDTDTMGRFCLPNSATVLIQOTVHAHRSDALQLGLGKHNYCRPNDRRRPWCYVQVGLK	140		
Qy	121	PLVOECNVHDCADGK	135		
Db	141	PLVOECNVHDCADGK	155		

```
RESULT 18
US-10-247-671-149
; Sequence 149, Application US/10247671
; Publication No. US20030194721A1
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; CURRENT FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 149
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: incyte ID No. US20030194721A1 1453334CD1
US-10-247-671-149

Query Match          99.7%; Score 791; DB 12; Length 431;
Best Local Similarity 99.3%; Pred. No. 1.4e-71;
Matches 134; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60
DB 21 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 80
QY 61 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 120
DB 81 KASDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPNDRRPPWCYVQVGLK 140
QY 121 PLVQECMVHDCADGK 135
DB 141 PLVQECMVHDCADGK 155

RESULT 19
US-10-407-821-2
; Sequence 2, Application US/10407821
; Publication No. US20030219386A1
; GENERAL INFORMATION:
; APPLICANT: IDELL, STEVEN
; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED
; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL
; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS
; FILE REFERENCE: UTSN:022US
; CURRENT APPLICATION NUMBER: US/10/407,821
; CURRENT FILING DATE: 2003-04-04
; PRIOR APPLICATION NUMBER: 60/414,202
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: 60/370,466
; PRIOR FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-407-821-2

Query Match          98.7%; Score 783; DB 12; Length 411;
Best Local Similarity 99.3%; Pred. No. 8.7e-71;
Matches 134; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFSNIHWCNCPKFGGQHCIEDKSKTCYEGNGHFYRG 60

RESULT 20
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match          64.3%; Score 510; DB 10; Length 88;
Best Local Similarity 100.0%; Pred. No. 5.2e-44;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 KTCYEGNGHFYRGKASTDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 107
DB 1 KTCYEGNGHFYRGKASTDTMTGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNCRPN 60
QY 108 RRRPPWCYVQVGLKPLVQECMVHDCADGK 135
DB 61 RRRPPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 21
US-09-880-503-9
; Sequence 9, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-9

Query Match          64.3%; Score 510; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 5.7e-44;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPON 107
Db 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPON 60

QY 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 22
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.218001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALAKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVYFKA 160

QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 23
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.219001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1

QY 48 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPON 107
Db 1 KTCYEGNGHYRGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPON 60

QY 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135
Db 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 22
US-09-987-457-18
; Sequence 18, Application US/09987457
; Publication No. US20030013150A1
; GENERAL INFORMATION:
; APPLICANT: Manosroi, Aranya
; APPLICANT: Manosroi, Jiradej
; APPLICANT: Tayapiwatana, Chatchai
; APPLICANT: Goetz, Friedrich
; APPLICANT: Werner, Rolf-Guenther
; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes
; FILE REFERENCE: 0652.218001
; CURRENT APPLICATION NUMBER: US/09/987,457
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,573
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 00 27 782.2
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; TYPE: PRT
; ORGANISM: Homo sapiens (tpa)
US-09-987-457-18

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALAKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVYFKA 160

QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 23
US-09-987-455-19
; Sequence 19, Application US/09987455
; Publication No. US20030049729A1
; GENERAL INFORMATION:
; APPLICANT: Aranya Manosroi
; APPLICANT: Jiradej Manosroi
; APPLICANT: Chatchai Tayapiwatana
; APPLICANT: Friedrich Goetz
; APPLICANT: Rolf-Guenther Werner
; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant
; FILE REFERENCE: 0652.219001
; CURRENT APPLICATION NUMBER: US/09/987,455
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/268,574
; PRIOR FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: GB 0027779.8
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 19
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-987-455-19

Query Match 41.4%; Score 328.5; DB 11; Length 527;
Best Local Similarity 46.3%; Pred. No. 6.7e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHY 58
Db 42 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 100

QY 59 RGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db 101 RGTWSTAESGAECTNNWSSALAKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVYFKA 160

QY 119 LKPLVQECMVHDCADG 134
Db 161 GKYSEFCSTPACSEG 176

RESULT 24
US-09-969-271-7
; Sequence 7, Application US/09969271
; Patent No. US20020098179A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc. (All designated States except GB and EP (GB));
; APPLICANT: Pfizer Limited (GB and EP (GB) only)
; TITLE OF INVENTION: Pharmaceutical Combinations
; FILE REFERENCE: PCS10951APME
; CURRENT APPLICATION NUMBER: US/09/969,271
; CURRENT FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: GB 0025473.0
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-969-271-7

Query Match 41.4%; Score 328.5; DB 9; Length 562;
Best Local Similarity 46.3%; Pred. No. 7.2e-25;
Matches 63; Conservative 14; Mismatches 54; Indels 5; Gaps 3;

QY 3 ELHQVP-SNCD---CLNGGTCVSNKYFSNIHWCNCPKFGQHCEDKSKTCYEGNGHY 58
Db 77 QCHSVPVKSCSEPRCFNGGTCQQAALYFSDP-VQCPEGFAGKCEIDTRATCYEDQGISY 135

QY 59 RGKASTDTMGRPCLPWNSATVLQOYTHAHRSDALQLGLGKHNYCRNPDRRRPWCYVQVG 118
Db 136 RGTWSTAESGAECTNNWSSALAKPYSGRRPDAILGLGNHNYCRNPDRDRSKWCYVYFKA 195

QY 119 LKPLVQECMVHDCADG 134
Db 196 GKYSEFCSTPACSEG 211

RESULT 25
US-09-974-298-145
; Sequence 145, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
```



```
; REFERENCE/DOCKET NUMBER: ANG-001/DIV (80144-0007)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 955-8787
; TELEFAX: (202) 955-3751
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Disulfide-bond
; LOCATION: group(11..19, 13..31, 33..42)
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-349-543--5

Query Match          36.9%; Score 293; DB 12; Length 49;
Best Local Similarity 100.0%; Pred. No. 2e-22;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKGGQHCEIDSKT 49
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKGGQHCEIDSKT 49

RESULT 29
US-09-880-503-7
; Sequence 7, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-7

Query Match          36.3%; Score 288; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 4.8e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKGGQHCEIDSK 48
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKGGQHCEIDSK 48

RESULT 30
US-09-880-503-2
; Sequence 2, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
```

```
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-2

Query Match          35.7%; Score 283; DB 10; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.9e-21;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKGGQHCEIDKS 47
Db 1 SNELHQVPSNCDCLNGGTCVSNKYFNSNIHWCNCPKFKGGQHCEIDKS 47

Search completed: December 3, 2003, 15:05:44
Job time : 14.8597 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:35:43 ; Search time 4.55172 Seconds
(without alignments)
818.010 Million cell updates/sec

Title: US-09-880-503-1

Perfect score: 510

Sequence: 1 KTCYEGNGHFGKASTDTM.....QVGLKPLVQRCMVHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep:*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep:*

5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	138	2	US-08-797-689-12
2	510	100.0	200	4	US-09-101-272G-73
3	510	100.0	208	4	US-09-101-272G-98
4	510	100.0	365	1	US-08-093-741-83
5	510	100.0	365	1	US-08-720-012-83
6	510	100.0	393	2	US-08-560-098A-44
7	510	100.0	393	3	US-08-967-024C-24
8	510	100.0	393	3	US-08-967-024C-25
9	510	100.0	411	1	US-08-087-163-1
10	510	100.0	411	1	US-08-286-748B-18
11	510	100.0	411	1	US-08-153-799-18
12	510	100.0	430	1	US-07-942-157A-3
13	510	100.0	430	6	5219569-2
14	510	100.0	431	4	US-09-101-272G-1
15	510	100.0	431	6	518829-1
16	510	100.0	432	2	US-08-560-098A-47
17	505	99.0	194	4	US-09-101-272G-80
18	505	99.0	201	4	US-09-101-272G-96
19	500	98.0	411	3	US-08-181-816-1
20	499	97.8	411	2	US-08-560-098A-48
21	489	95.9	89	4	US-09-101-272G-62
22	486	95.3	157	3	US-08-142-590B-25
23	241	47.3	477	2	US-08-560-098A-51
24	226	44.3	527	1	US-07-609-510B-16
25	226	44.3	527	2	US-08-811-949-39
26	226	44.3	527	5	PCT-US91-01025A-2
27	226	44.3	527	6	5185259-8

28	226	44.3	527	6	5520913-1
29	226	44.3	546	6	5200340-6
30	226	44.3	562	2	US-08-811-949-43
31	226	44.3	562	2	US-08-560-098A-50
32	226	44.3	562	2	US-08-883-795A-38
33	226	44.3	562	6	5185259-3
34	226	44.3	562	6	5200340-2
35	226	44.3	562	6	5344773-2
36	221	43.3	437	2	US-08-811-949-49
37	221	43.3	437	2	US-08-811-949-51
38	221	43.3	437	2	US-08-811-949-55
39	221	43.3	437	2	US-08-811-949-57
40	218	42.7	356	1	US-08-427-640-8
41	218	42.7	562	6	5244676-5
42	217	42.5	378	4	US-09-553-498-10
43	217	42.5	378	4	US-09-618-869-10
44	216	42.4	472	2	US-08-811-949-63
45	215	42.2	326	4	US-09-411-977-3

ALIGNMENTS

RESULT 1

US-08-797-689-12
; Sequence 12, Application US/08797689
; Patent No. 5875969
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; APPLICANT: Fournier, Alain
; APPLICANT: Guitton, Jean-Dominique
; APPLICANT: Jung, Gerard
; APPLICANT: Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; TITLE OF INVENTION: PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08797,689
; FILING DATE: 31-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear


```
; MOLECULE TYPE: protein
US-08-797-689-12

Query Match      100.0%; Score 510; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 8.5e-54;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 110

Qy 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPCWYVQVGLKPLVQECMVHDCADGK 138

RESULT 2
US-09-101-272G-73
; Sequence 73, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 73
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: ATF domain of uPA
US-09-101-272G-73

Query Match      100.0%; Score 510; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPCWYVQVGLKPLVQECMVHDCADGK 155

RESULT 3
US-09-101-272G-98
; Sequence 98, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ATTHI-ML chimeric protein
US-09-101-272G-98

; MOLECULE TYPE: protein
US-08-797-689-12

Query Match      100.0%; Score 510; DB 4; Length 208;
Best Local Similarity 100.0%; Pred. No. 1.4e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 49 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 108

Qy 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88
Db 109 RRRPCWYVQVGLKPLVQECMVHDCADGK 136

RESULT 4
US-08-093-741-83
; Sequence 83, Application US/08093741
; Patent No. 5681721
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, GERD J.
; APPLICANT: WENDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, Mckeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/093.741
; FILING DATE: 20-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-093-741-83

Query Match      100.0%; Score 510; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.7e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61

Qy 61 RRRPCWYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPCWYVQVGLKPLVQECMVHDCADGK 89
```

RESULT 5
US-08-720-012-83
; Sequence 83, Application US/08720012
; Patent No. 5747291
; GENERAL INFORMATION:
; APPLICANT: STEFFENS, Gerd J.
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: SCHNEIDER, JOHANNES
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: SAUNDERS, DEREK J.
; TITLE OF INVENTION: BIFUNCTIONAL UROKINASE VARIANTS WITH
; TITLE OF INVENTION: IMPROVED FIBRINOLYTIC CHARACTERISTICS AND THROMBIN
; TITLE OF INVENTION: INHIBITING EFFECT
; NUMBER OF SEQUENCES: 83
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N. W. Suite 700
; CITY: Washington, D.C.
; COUNTRY: U.S.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720.012
; FILING DATE: 27-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,741
; FILING DATE: 20-JUL-1993
; APPLICATION NUMBER: DE P43 23 754.1
; FILING DATE: 15-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/41345
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)628-8800
; TELEFAX: (202)628-8844
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-720-012-83
Query Match 100.0%; Score 510; DB 1; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.7e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 61
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 6
US-08-560-098A-44
; Sequence 44, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: HEINZEL-WIELAND, REGINA
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560.098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-44
Query Match 100.0%; Score 510; DB 2; Length 393;
Best Local Similarity 100.0%; Pred. No. 2.9e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 3 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 62
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 63 RRRPWCYVQVGLKPLVQECMVHDCADGK 90

RESULT 7
US-08-967-024C-24
; Sequence 24, Application US/08967024C
; Patent No. 6133011
; GENERAL INFORMATION:
; APPLICANT: WENNDT, STEPHAN
; APPLICANT: STEFFENS, Gerd Josef
; APPLICANT: JANOSCH, Elke
; APPLICANT: HEINZEL-WIELAND, REGINA
; TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967.024C
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

06:30:55 2003

AN NUMBER: P 44 42 665.8
DE: 30-NOV-1994
ENT INFORMATION:
ANS, Joseph D.
ION NUMBER: 26, 269
E/DOCKET NUMBER: 148/42444
ICATION INFORMATION:
DE: (202) 628-8800
RSE: (202) 628-8844
CHARACTERISTICS: 24:
393 amino acids
amino acid
EDNESS:
GY: linear
TYPE: protein
4C-24

h 100.0%; Score 510; DB 3; Length 393;
1 Similarity 100.0%; Pred. No. 2.9e-53;
88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLTQTYHAHRSALQGLGKHYCRNPDN 60
3 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLTQTYHAHRSALQGLGKHYCRNPDN 62

61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90

57-024C-25
US 25 Application US/08967024C
No. 61301
RA INFORMATION:
APPLICANT: WENDET, Stephan
APPLICANT: STEFFENS, Gerd Josef
APPLICANT: JANOSCH, Elke
APPLICANT: HEINZEL-WIELAND, Regina
TITLE OF INVENTION: Chimeric Proteins having Fibrinolytic
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/967,024C
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 1944
EPO NUMBER: 1944
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26, 269
REFERENCE/DOCKET NUMBER: 148/42444
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 393 amino acids
TYPE: amino acid
STRANDEDNESS:

us-09-880-503-1.rai

MOLECULE TYPE: protein
US-08-967-024C-25

Query Match 100.0%; Score 510; DB 3; Length 393;
Best Local Similarity 100.0%; Pred. No. 2.9e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLTQTYHAHRSALQGLGKHYCRNPDN 60
DB 3 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLTQTYHAHRSALQGLGKHYCRNPDN 62

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
DB 63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90

RESULT 9
US-08-087-163-1
Sequence 1, Application US/08087163
Patent No. 5472692
GENERAL INFORMATION:
APPLICANT: Liu, Jian-Ning
APPLICANT: Gurewicz, Victor
TITLE OF INVENTION: PRO-UKINASE MUTANTS
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 502 or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: Patent in Release #1.0, Version 5.11
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,163
FILING DATE: 07/02/93
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04353/003001
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 411
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A

US-08-087-163-1

Query Match 100.0%; Score 510; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.1e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLTQTYHAHRSALQGLGKHYCRNPDN 60
DB 48 KTCYEGNGHYRGKASTDTMGRCPLPNSATVLTQTYHAHRSALQGLGKHYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGK 135

RESULT 10

US-08-286-748B-18
; Sequence 18, Application US/08286748B
; Patent No. 5759542
; GENERAL INFORMATION:
; APPLICANT: Victor Gurewlich
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DELIVERY
; TITLE OF INVENTION: OF DRUGS BY PLATELETS FOR THE TREATMENT OF
; TITLE OF INVENTION: CARDIOVASCULAR AND OTHER DISEASES
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286,748B
; FILING DATE: August 5, 1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: J. Peter Fasse
; REGISTRATION NUMBER: 32,983
; REFERENCE/DOCKET NUMBER: 04547/013001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-286-748B-18

Query Match 100.0%; Score 510; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.1e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPCYVQVGLKPLVQECMVHDCADGK 135
RESULT 11
US-08-153-799-18
; Sequence 18, Application US/08153799
; Patent No. 5766883
; GENERAL INFORMATION:
; APPLICANT: Ballance, David J
; APPLICANT: Goodey, Andrew R
; TITLE OF INVENTION: Polypeptides
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: R Hain Swope, BOC Health Care Inc
; STREET: 100 Mountain Avenue
; CITY: Murray Hill
; STATE: New Jersey
; COUNTRY: USA

ZIP: 07974

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/153,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847975
; FILING DATE: 06-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 8909916.2
; FILING DATE: 29-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB90/00650
; FILING DATE: 26-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/775952
; FILING DATE: 29-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Swope, R Hain
; REGISTRATION NUMBER: 24864
; REFERENCE/DOCKET NUMBER: 92H832
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 665 2400
; TELEFAX: (908) 771 6159
; TELEX: 219484
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 411 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-153-799-18

Query Match 100.0%; Score 510; DB 1; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.1e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRRPCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPCYVQVGLKPLVQECMVHDCADGK 135

RESULT 12

US-07-942-157A-3
; Sequence 3, Application US/07942157A
; Patent No. 5648253
; GENERAL INFORMATION:
; APPLICANT: Wei, Cha-Mer
; TITLE OF INVENTION: Inhibitor-Resistant Urokinase
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/942,157A

```

; FILING DATE: 19920908
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631673
; FILING DATE: 20-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: TS1108Cont.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)815-6508
; TELEFAX: (404)815-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 430 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; OTHER INFORMATION: /label= peptide
; OTHER INFORMATION: /note= "WAP signal"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 198..203
; OTHER INFORMATION: /label= modified
; OTHER INFORMATION: /note= "six amino acids deleted in mutant"
; US-07-942-157A-3

Query Match 100.0%; Score 510; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 67 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 126

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 127 RRRPWCYVQVGLKPLVQECMVHDCADGK 154

RESULT 13
5219569-2
; Patent No. 5219569
; APPLICANT: BLABER, MICHAEL; HEYNEKER, HERBERT L.; VEHAR,
; GORDON A.
; TITLE OF INVENTION: PROTEASE RESISTANT UROKINASE
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; FILING DATE: 16-AUG-1985
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 725,468
; FILING DATE: 22-APR-1985
; SEQ ID NO: 2:
; LENGTH: 430
5219569-2

Query Match 100.0%; Score 510; DB 6; Length 430;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 14
US-09-101-272G-1
; Sequence 1, Application US/09101272G
; Patent No. 6509445
; GENERAL INFORMATION:
; APPLICANT: Nissin Food Products Co., Ltd.
; TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
; FILE REFERENCE: Q50979
; CURRENT APPLICATION NUMBER: US/09/101,272G
; CURRENT FILING DATE: 1998-07-08
; PRIOR APPLICATION NUMBER: JP 1059/1996
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1:
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (21)..()
; OTHER INFORMATION:
; NAME/KEY: misc_feature
; LOCATION: (20)..()
; OTHER INFORMATION: Urokinase-type plasminogen activator (uPA)
; US-09-101-272G-1

Query Match 100.0%; Score 510; DB 4; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 15
5188829-1
; Patent No. 5188829
; APPLICANT: KOBAYASHI, YO-ICHI; OMORI, MUNEKI; YAMADA, CHIKAKO
; TITLE OF INVENTION: RAPIDLY ACTING PROUROKINASE
; NUMBER OF SEQUENCES: 23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/340,007
; FILING DATE: 18-AUG-1988
; SEQ ID NO: 1:
; LENGTH: 431
5188829-1

Query Match 100.0%; Score 510; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWN SATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 16
US-08-560-098A-47
; Sequence 47, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WENNDT, Stephan
```

APPLICANT: HEINZEL-WIELAND, Regina
APPLICANT: STEFFENS, Gerd Josef
TITLE OF INVENTION: Proteins having Fibrinolytic and
TITLE OF INVENTION: Coagulation-inhibiting Properties
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
STREET: 1200 G Street, N.W., Suite 700
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,098A
FILING DATE: 17-NOV-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P 44 40 892.7
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: EVANS, Joseph D.
REGISTRATION NUMBER: 26,269
REFERENCE/DOCKET NUMBER: 148/42448
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-8800
TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-560-098A-47

Query Match 100.0%; Score 510; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.3e-53;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60
Db 69 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 128
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 129 RRPWCYVQVGLKPLVQECMVHDCADGK 156

RESULT 17
US-09-101-272G-80
Sequence 80, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 80
LENGTH: 194
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATFHI chimeric protein
US-09-101-272G-80

Query Match 99.0%; Score 505; DB 4; Length 194;
Best Local Similarity 100.0%; Pred. No. 5.1e-53;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60
Db 49 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 108
QY 61 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 109 RRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 18
US-09-101-272G-96
Sequence 96, Application US/09101272G
Patent No. 6509445
GENERAL INFORMATION:
APPLICANT: Nissin Food Products Co., Ltd.
TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
FILE REFERENCE: Q50979
CURRENT APPLICATION NUMBER: US/09/101,272G
CURRENT FILING DATE: 1998-07-08
PRIOR APPLICATION NUMBER: JP 1059/1996
PRIOR FILING DATE: 1996-01-08
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn version 3.1
SEQ ID NO 96
LENGTH: 201
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: ATFHI-CL chimeric protein
US-09-101-272G-96

Query Match 99.0%; Score 505; DB 4; Length 201;
Best Local Similarity 100.0%; Pred. No. 5.3e-53;
Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 60
Db 49 KTCYEGNGHFYRGKASTDTMGRLPWN SATVLQQTYYHAHRS DALQLGLGKHNYCRNPDN 108
QY 61 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 109 RRPWCYVQVGLKPLVQECMVHDCADG 135

RESULT 19
US-09-181-816-1
Sequence 1, Application US/09181816
Patent No. 6277818
GENERAL INFORMATION:
APPLICANT: MAZAR, Andrew P.
APPLICANT: JONES, Terence R.
TITLE OF INVENTION: CYCLIC PEPTIDE LIGANDS THAT TARGET UROKINASE
TITLE OF INVENTION: PLASMINOGEN ACTIVATOR RECEPTOR
FILE REFERENCE: 329042000300 SIDN 1-7
CURRENT APPLICATION NUMBER: US/09/181,816
CURRENT FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-181-816-1

Query Match 98.0%; Score 500; DB 3; Length 411;
Best Local Similarity 98.9%; Pred. No. 4.9e-52;
Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPDN 60
 4B KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPDN 107
 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
 108 RRRPWCYVQVGLKPLVQECWVHDCADGK 135

20
 60-098A-48
 nce 48, Application US/08560098A
 at No. 5976841

REAL INFORMATION:
 APPLICANT: WENNDT, Stephan
 REPLICANT: HEINZEL-WIELAND, Regina
 REPLICANT: STEFFENS, Gerd Josef
 TITLE OF INVENTION: Proteins having fibrinolytic and
 title of invention: Coagulation-inhibiting Properties
 NUMBER OF SEQUENCES: 60
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Evensen, McKeown, Edwards & Lenahan
 STREET: 1200 G Street, N.W., Suite 700
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
 CURRENT APPLICATION NUMBER: US/08/560,098A
 FILING DATE: 17-NOV-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P 44 40 892.7
 FILING DATE: 17-NOV-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: EVANS, Joseph D.
 REGISTRATION NUMBER: 26,269
 REFERENCE/DOCKET NUMBER: 148/42448
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-8800
 TELEFAX: (202) 628-8844
 INFORMATION FOR SEQ ID NO: 48:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 411 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-560-098A-48

Query Match 97.8%; Score 499; DB 2; Length 411;
 Best Local Similarity 98.9%; Pred. No. 6.5e-52;
 Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPDN 60
 DB 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPDN 107
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
 DB 108 RRRPWCYVQVGLKPLVQECWVHDCADGK 135

RESULT 21
 US-09-101-272G-62
 ; Sequence 62, Application US/09101272G
 ; Patent No. 6509445
 ; GENERAL INFORMATION:
 ; APPLICANT: Nissin Food Products Co., Ltd.

TITLE OF INVENTION: CANCEROUS METASTASIS INHIBITOR
 FILE REFERENCE: Q50979
 CURRENT APPLICATION NUMBER: US/09/101,272G
 CURRENT FILING DATE: 1998-07-08
 PRIOR APPLICATION NUMBER: JP 1059/1996
 PRIOR FILING DATE: 1996-01-08
 NUMBER OF SEQ ID NOS: 107
 SOFTWARE: Patent in version 3.1
 SEQ ID NO 62
 LENGTH: 89
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc feature
 OTHER INFORMATION: residues 43-131 of the ATF domain of uPA
 US-09-101-272G-62

Query Match 95.9%; Score 489; DB 4; Length 89;
 Best Local Similarity 100.0%; Pred. No. 1.7e-51;
 Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPDN 60
 DB 6 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQGLGKHNCRNPDN 65
 QY 61 RRRPWCYVQVGLKPLVQECWVHDC 84
 DB 66 RRRPWCYVQVGLKPLVQECWVHDC 89

RESULT 22
 US-08-142-590B-25
 ; Sequence 25, Application US/08142590B
 ; Patent No. 6120765
 ; GENERAL INFORMATION:
 ; APPLICANT: HIBINO, Tashihiko; TAKAHASHI, Tadahito; HORII, Izumi; and
 ; TITLE OF INVENTION: UROKINASE PLASMINOGEN ACTIVATOR FRAGMENTS
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: LAHIVE & COCKFIELD, LLP
 ; STREET: 28 State Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/142,590B
 ; FILING DATE: 25-OCT-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/042,318
 ; FILING DATE: 02-APR-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Myers, Paul L.
 ; REGISTRATION NUMBER: 35,965
 ; REFERENCE/DOCKET NUMBER: MGP-009CP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 227-7400
 ; TELEFAX: (617) 227-5941
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 157 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; FRAGMENT TYPE: internal
 ; US-08-142-590B-25

Query Match 95.3%; Score 486; DB 3; Length 157;

```
; Best Local Similarity 95.5%; Pred. No. 7.5e-51;
Matches 84; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFYRGKASTDTMGRPCLPWNSAYVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 23
US-08-560-098A-51
; Sequence 51, Application US/08560098A
; Patent No. 5976841
; GENERAL INFORMATION:
; APPLICANT: WNEU, Stephan
; APPLICANT: HEINZEL-WIELAND, Regina
; APPLICANT: STEFFENS, Gerd Josef
; TITLE OF INVENTION: Proteins having Fibrinolytic and
; TITLE OF INVENTION: Coagulation-inhibiting Properties
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Evenson, McKeown, Edwards & Lenahan
; STREET: 1200 G Street, N.W., Suite 700
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/560,098A
; FILING DATE: 17-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P 44 40 892.7
; FILING DATE: 17-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: EVANS, Joseph D.
; REGISTRATION NUMBER: 26,269
; REFERENCE/DOCKET NUMBER: 148/42448
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-8800
; TELEFAX: (202) 628-8844
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-560-098A-51

Query Match 47.3%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 7.4e-21;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 61
Db 127 TCYKQGVYIRGTWSTSESGAQCINWNSLLTFRTYNGRSDAITLGLGNHNYCRNPDN 186
QY 62 RRPWCYVQVGLKPLVQECMVHDCA 85
Db 187 SKPCVCYKASKFILEFCVPVCS 210

RESULT 24
US-07-609-510B-16
; Sequence 16, Application US/07609510B
```

```
; Patent No. 5326700
; GENERAL INFORMATION:
; APPLICANT: Berg et al.
; TITLE OF INVENTION: Method for Altering Post-Translational Processing of Tissue F
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN.
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,510B
; FILING DATE: 19901106
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 527 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-510B-16

Query Match 44.3%; Score 226; DB 1; Length 527;
Best Local Similarity 47.7%; Pred. No. 5.3e-19;
Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 61
Db 91 TCYEDQGISYRGTWSTAESGAECTNWSALAKPKYSGRRPDAIRLGLGNHNYCRNPDN 150
QY 62 RRPWCYVQVGLKPLVQECMVHDCADG 87
Db 151 SKPCVCYKASKFILEFCVPVCS 176

RESULT 25
US-08-811-949-39
; Sequence 39, Application US/08811949
; Patent No. 5840533
; GENERAL INFORMATION:
; APPLICANT: NIWA, MINEO
; APPLICANT: SAITO, YOSHIMASA
; APPLICANT: SASAKI, HITOSHI
; APPLICANT: HAYASHI, MASAKO
; APPLICANT: NOTANI, JOUJI
; APPLICANT: KOBAYASHI, MASAKAZU
; TITLE OF INVENTION: TISSUE PLASMINOGEN ACTIVATOR
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,949
; FILING DATE: 05-MAR-1997
; CLASSIFICATION: 435
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:43:04 ; Search time 9.03448 Seconds
(without alignments)
1811.566 Million cell updates/sec

Title: US-09-880-503-1
Perfect score: 510
Sequence: 1 KTCYEGNGHYRGKASTDTM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	510	100.0	88	US-09-880-503-1	Sequence 1, Appli
2	510	100.0	96	US-09-880-503-9	Sequence 9, Appli
3	510	100.0	135	US-09-880-503-4	Sequence 4, Appli
4	510	100.0	138	US-09-884-186-12	Sequence 12, Appli
5	510	100.0	138	US-10-237-667-12	Sequence 12, Appli
6	510	100.0	138	US-10-237-708-12	Sequence 12, Appli
7	510	100.0	138	US-10-237-866-12	Sequence 12, Appli
8	510	100.0	138	US-10-237-871-12	Sequence 12, Appli
9	510	100.0	138	US-10-237-624-12	Sequence 12, Appli
10	510	100.0	143	US-09-880-503-8	Sequence 8, Appli
11	510	100.0	337	US-10-106-698-6266	Sequence 6266, Ap
12	510	100.0	403	US-09-880-503-6	Sequence 6, Appli
13	510	100.0	411	US-09-880-503-3	Sequence 3, Appli
14	510	100.0	431	US-09-264-468B-1	Sequence 1, Appli
15	510	100.0	431	US-10-301-822-161	Sequence 161, App

16	510	100.0	431	12	US-10-131-985-21	Sequence 21, Appli
17	510	100.0	431	14	US-10-076-421-2	Sequence 2, Appli
18	510	100.0	431	15	US-10-171-311-184	Sequence 184, App
19	510	100.0	431	15	US-10-193-656-4	Sequence 4, Appli
20	508	99.6	431	12	US-10-247-671-149	Sequence 149, App
21	500	98.0	411	12	US-10-407-821-2	Sequence 2, Appli
22	226	44.3	527	11	US-09-987-457-18	Sequence 18, Appli
23	226	44.3	527	11	US-09-987-455-19	Sequence 19, Appli
24	226	44.3	562	9	US-09-969-271-7	Sequence 7, Appli
25	226	44.3	562	10	US-09-974-298-145	Sequence 145, App
26	226	44.3	562	12	US-10-443-701-4	Sequence 4, Appli
27	226	44.3	562	15	US-10-193-656-8	Sequence 8, Appli
28	215	42.2	326	14	US-10-057-951-3	Sequence 3, Appli
29	215	42.2	354	11	US-09-987-457-10	Sequence 10, Appli
30	215	42.2	354	11	US-09-987-455-11	Sequence 11, Appli
31	215	42.2	372	9	US-09-084-491A-3	Sequence 3, Appli
32	215	42.2	372	14	US-10-102-704-3	Sequence 3, Appli
33	215	42.2	377	11	US-09-987-455-8	Sequence 8, Appli
34	193.5	37.9	560	10	US-09-912-559-3	Sequence 3, Appli
35	193.5	37.9	560	10	US-09-912-559-4	Sequence 4, Appli
36	193.5	37.9	560	15	US-10-172-712-32	Sequence 32, Appli
37	193	37.8	655	15	US-10-172-712-28	Sequence 28, Appli
38	176	34.5	343	11	US-09-987-457-14	Sequence 14, Appli
39	176	34.5	343	11	US-09-987-455-15	Sequence 15, Appli
40	175	34.3	339	11	US-09-987-457-12	Sequence 12, Appli
41	175	34.3	339	11	US-09-987-455-13	Sequence 13, Appli
42	173	33.9	83	12	US-09-981-151A-86	Sequence 86, Appli
43	172	33.7	331	11	US-09-987-457-11	Sequence 11, Appli
44	172	33.7	331	11	US-09-987-455-12	Sequence 12, Appli
45	171	33.5	79	12	US-09-981-151A-85	Sequence 85, Appli

ALIGNMENTS

RESULT 1
US-09-880-503-1
; Sequence 1, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880, 503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212, 847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-1

Query Match	100.0%	Score	510;	DB	10;	Length	88;
Best Local Similarity	100.0%	Pred. No.	3.5e-52;				
Matches	88;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	KTCYEGNGHYRGKASTDTMGRPCLPWSNGATVLCQTYHAHRSALQGLGKHNYCRNPDN	60				
Db	1	KTCYEGNGHYRGKASTDTMGRPCLPWSNGATVLCQTYHAHRSALQGLGKHNYCRNPDN	60				
QY	61	RRRPWCYVQVGLKPLVQECMVHDCADGK	88				
Db	61	RRRPWCYVQVGLKPLVQECMVHDCADGK	88				

RESULT 2
US-09-880-503-9
; Sequence 9, Application US/09880503

```
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-9

Query Match      100.0%; Score 510; DB 10; Length 96;
Best Local Similarity 100.0%; Pred. No. 3.9e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

RESULT 3
US-09-880-503-4
; Sequence 4, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Root
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-880-503-4

Query Match      100.0%; Score 510; DB 10; Length 135;
Best Local Similarity 100.0%; Pred. No. 5.6e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      48 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 107

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 4
US-09-880-186-12
; Sequence 12, Application US/09984186
; Patent No. US20020151011A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/984,186
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-984-186-12

Query Match      100.0%; Score 510; DB 10; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.7e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db      51 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 110

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 5
US-10-237-667-12
; Sequence 12, Application US/10237667
; Publication No. US20030022308A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guitton, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
```



```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,866
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-866-12
Query Match 100.0%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 5,7e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 110

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 8
US-10-237-871-12
; Sequence 12, Application US/10237871
; Publication No. US20030036172A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
;
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,624
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,871
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
;
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-871-12
Query Match 100.0%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 5,7e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFRGKASTDTMGRCPLPWNATVLOQTYHAHRS DALQLGLGKHNYCRNPDN 110

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 9
US-10-237-624-12
; Sequence 12, Application US/10237624
; Publication No. US20030082747A1
; GENERAL INFORMATION:
; APPLICANT: Fleer, Reinhard
; Fournier, Alain
; Guittion, Jean-Dominique
; Jung, Gerard
; Yeh, Patrice
;
; TITLE OF INVENTION: NOVEL BIOLOGICALLY ACTIVE POLYPEPTIDES,
; PREPARATION THEREOF AND PHARMACEUTICAL COMPOSITION
; CONTAINING SAID POLYPEPTIDES
;
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: System 7.1
; SOFTWARE: Word 5.1 (PatentIn)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/237,624
; FILING DATE: 10-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:

```

```
; APPLICATION NUMBER: US/08/797,689
; FILING DATE: 31-JAN-1997
; APPLICATION NUMBER: US 08/256,927
; FILING DATE: 28-JUL-1994
; APPLICATION NUMBER: FR 92/01064
; FILING DATE: 31-JAN-1992
; APPLICATION NUMBER: PCT/FR93/00085
; FILING DATE: 28-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith Ph.D., Julie K.
; REGISTRATION NUMBER: P-38,619
; REFERENCE/DOCKET NUMBER: ST92006-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3829
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 138 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-237-624-12

Query Match 100.0%; Score 510; DB 15; Length 138;
Best Local Similarity 100.0%; Pred. No. 5.7e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 51 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 110

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 111 RRRPWCYVQVGLKPLVQECMVHDCADGK 138

RESULT 10
US-09-880-503-8
; Sequence 8, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-8

Query Match 100.0%; Score 510; DB 10; Length 143;
Best Local Similarity 100.0%; Pred. No. 5.9e-52;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 11
US-09-880-503-3
```

```
US-10-106-698-6266
; Sequence 6266, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypepti
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6266
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6266

Query Match 100.0%; Score 510; DB 15; Length 337;
Best Local Similarity 100.0%; Pred. No. 1.5e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 74 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 133

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 12
US-09-880-503-6
; Sequence 6, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TITLE OF INVENTION: TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-6

Query Match 100.0%; Score 510; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 1.8e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
Db 48 KTCYEGNGHFGYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 13
US-09-880-503-3
```

```
; Sequence 3, Application US/09880503
; Patent No. US20020131964A1
; GENERAL INFORMATION:
; APPLICANT: CINES, Douglas B
; APPLICANT: HIGAZI, Abd Al-Roof
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING MUSCLE CELL AND
; TISSUE CONTRACTABILITY
; FILE REFERENCE: 9596-331
; CURRENT APPLICATION NUMBER: US/09/880,503
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/212,847
; PRIOR FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-503-3

Query Match      100.0%; Score 510; DB 10; Length 411;
Best Local Similarity 100.0%; Pred. No. 1.9e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 60
Db      48 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 107

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 14
US-09-264-468B-1
; Sequence 1, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Jieyi
; APPLICANT: Nienaber, Vicki L.
; APPLICANT: Henkin, Jack
; APPLICANT: Smith, Richard A.
; APPLICANT: Walter, Karl A.
; APPLICANT: Severin, Jean M.
; APPLICANT: Edalji, Rohinton
; APPLICANT: Johnson Jr., Robert W.
; APPLICANT: Holzman, Thomas F.
; TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
; FILE REFERENCE: 6310-US.P1
; CURRENT APPLICATION NUMBER: US/09/264,468B
; CURRENT FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: US 09/036,361
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(20)
; OTHER INFORMATION: Leader sequence
; NAME/KEY: VARIANT
; LOCATION: (279)...(279)
; OTHER INFORMATION: Xaa = any amino acid
; NAME/KEY: VARIANT
; LOCATION: (302)...(302)
; OTHER INFORMATION: Xaa = any amino acid
US-09-264-468B-1

Query Match      100.0%; Score 510; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 60
Db      48 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 107

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

US-09-880-503-1
; Sequence 1, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-0292P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match      100.0%; Score 510; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 60
Db      68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 127

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 15
US-10-301-822-161
; Sequence 161, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: BURGART, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; THERAPY OF COLON CANCER
; FILE REFERENCE: MEM01-0292P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 161
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-161

Query Match      100.0%; Score 510; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 60
Db      68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQO TYHAHRS DALQLGLGKHNYCRNPDN 127

Qy      61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db      128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 16
US-10-131-985-21
; Sequence 21, Application US/10131985
; Publication No. US20030199440A1
; GENERAL INFORMATION:
; APPLICANT: Dack, Kevin N
; APPLICANT: Davies, Michael J
; APPLICANT: Fish, Paul V
; APPLICANT: Huggins, Jonathan P
; APPLICANT: McIntosh, Fraser S
; APPLICANT: Occleston, Nicholas L
; TITLE OF INVENTION: Composition
; FILE REFERENCE: PCS 10391A
; CURRENT APPLICATION NUMBER: US/10/131,985
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: US/09/726,295
```

```
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: GB 9930768.8
; PRIOR FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-985-21

Query Match      100.0%; Score 510; DB 12; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 17
US-10-076-421-2
; Sequence 2, Application US/10076421
; Publication No. US20020193304A1
; GENERAL INFORMATION:
; APPLICANT: WADA, MANABU
; APPLICANT: WADA, NAKO
; TITLE OF INVENTION: ANTI-HIV AGENTS
; FILE REFERENCE: HAYAK-9
; CURRENT APPLICATION NUMBER: US/10/076,421
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: JP 2001-42655
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: JP 2001-184284
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-421-2

Query Match      100.0%; Score 510; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 18
US-10-171-311-184
; Sequence 184, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
```

```
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 184
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-184

Query Match      100.0%; Score 510; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 19
US-10-193-656-4
; Sequence 4, Application US/10193656
; Publication No. US20030096733A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Rikard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/1J577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00749
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(431)
US-10-193-656-4

Query Match      100.0%; Score 510; DB 15; Length 431;
Best Local Similarity 100.0%; Pred. No. 2e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 60
Db 68 KTCYEGNGHFYRGKASTDTMGRPCLPWN SATVLQQTYYHAHRSDALQLGLGKHN YCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155
```

Db 128 RRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 20

US-10-247-671-149

; Sequence 149, Application US/10247671

; Publication No. US20030194721A1

; GENERAL INFORMATION:

; APPLICANT: Mikita, Thomas

; APPLICANT: Shiffman, Dov

; APPLICANT: Porter, Gordon, J.

; APPLICANT: Kaser, Matthew R.

; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS

; FILE REFERENCE: PA-0050 US

; CURRENT APPLICATION NUMBER: US/10/247,671

; PRIOR FILING DATE: 2002-09-18

; PRIOR APPLICATION NUMBER: 60/323,784

; PRIOR FILING DATE: 2001-09-19

; NUMBER OF SEQ ID NOS: 186

; SOFTWARE: PERL Program

; SEQ ID NO 149

; LENGTH: 431

; TYPE: PRT

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20030194721A1 14533334CD1

US-10-247-671-149

Query Match 99.6%; Score 508; DB 12; Length 431;

Best Local Similarity 98.9%; Pred. No. 3.4e-51;

Matches 87; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60

Db 68 KTCYEGNGHFRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 127

Qy 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88

Db 128 RRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 21

US-10-407-821-2

; Sequence 2, Application US/10407821

; Publication No. US20030219386A1

; GENERAL INFORMATION:

; APPLICANT: IDELL, STEVEN

; TITLE OF INVENTION: INTRAPLEURAL SINGLE-CHAIN UROKINASE ALONE OR COMPLEXED

; TITLE OF INVENTION: TO ITS SOLUBLE RECEPTOR PROTECTS AGAINST PLEURAL

; TITLE OF INVENTION: ADHESIONS IN TETRACYCLINE-INDUCED PLEURITIS IN RABBITS

; FILE REFERENCE: UTSN:022US

; CURRENT APPLICATION NUMBER: US/10/407,821

; PRIOR FILING DATE: 2003-04-04

; PRIOR APPLICATION NUMBER: 60/414,202

; PRIOR FILING DATE: 2002-09-27

; PRIOR APPLICATION NUMBER: 60/370,456

; PRIOR FILING DATE: 2002-04-05

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: PatentIn ver. 2.1

; SEQ ID NO 2

; LENGTH: 411

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-407-821-2

Query Match 98.0%; Score 500; DB 12; Length 411;

Best Local Similarity 98.9%; Pred. No. 2.8e-50;

Matches 87; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 60

Db 48 KTCYEGNGHFRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 107

Qy 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88

Db 108 RRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 22

US-09-987-457-18

; Sequence 18, Application US/09987457

; Publication No. US20030013150A1

; GENERAL INFORMATION:

; APPLICANT: Manosroi, Aranya

; APPLICANT: Manosroi, Jiradej

; APPLICANT: Tavapiwatana, Chatchai

; APPLICANT: Goetz, Friedrich

; APPLICANT: Werner, Rolf-Guenther

; TITLE OF INVENTION: Methods for Large Scale Protein Production in Prokaryotes

; FILE REFERENCE: 0652.2180001

; CURRENT APPLICATION NUMBER: US/09/987,457

; CURRENT FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/268,573

; PRIOR FILING DATE: 2001-02-15

; PRIOR APPLICATION NUMBER: GB 00 27 782.2

; PRIOR FILING DATE: 2000-11-14

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 18

; LENGTH: 527

; TYPE: PRT

; ORGANISM: Homo sapiens (tPA)

US-09-987-457-18

Query Match 44.3%; Score 226; DB 11; Length 527;

Best Local Similarity 47.7%; Pred. No. 4.3e-19;

Matches 41; Conservative 9; Mismatches 36; Indels 0; Gaps 0;

Qy 2 TCYEGNGHFRGKASTDTMGRPCLPWN SATVLQOQTYHAHRS DALQLGLGKHNYCRNPDN 61

Db 91 TCYEDQGISYRGTWSTAESGAECTNNSSALAAQKPYSGRRPDAIRLGLGNHNYCRNPD 150

Qy 62 RRPWCYVQVGLKPLVQECMVHDCADG 87

Db 151 SKPCYVFKAGYSSEFCSTPACSEG 176

RESULT 23

US-09-987-455-19

; Sequence 19, Application US/09987455

; Publication No. US20030049729A1

; GENERAL INFORMATION:

; APPLICANT: Aranya Manosroi

; APPLICANT: Jiradej Manosroi

; APPLICANT: Chatchai Tavapiwatana

; APPLICANT: Friedrich Goetz

; APPLICANT: Rolf-Guenther Werner

; TITLE OF INVENTION: Methods for Large Scale Production of Recombinant

; TITLE OF INVENTION: DNA-Derived tPA or K2S Molecules

; FILE REFERENCE: 0652.2190001

; CURRENT APPLICATION NUMBER: US/09/987,455

; CURRENT FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 60/268,574

; PRIOR FILING DATE: 2001-02-15

; PRIOR APPLICATION NUMBER: GB 0027779.8

; PRIOR FILING DATE: 2000-11-14

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 19

; LENGTH: 527

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-987-455-19

Query Match 44.3%; Score 226; DB 11; Length 527;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:58 ; Search time 4.55172 Seconds
(without alignments)
1859.261 Million cell updates/sec

Title: US-09-880-503-1
Perfect score: 510
Sequence: 1 KTCYEGNGHFRGKASTDTM.....QVGLKPLVQECMVHDCADGK 88

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76:*
1: Pir1:*
2: Pir2:*
3: Pir3:*
4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	510	100.0	431	1 UKHU	u-plasminogen acti
2	475	93.1	433	1 UKBAY	u-plasminogen acti
3	420	82.4	442	1 UKPG	u-plasminogen acti
4	406	79.6	432	1 S18932	u-plasminogen acti
5	396	77.6	433	1 JN0560	u-plasminogen acti
6	387	75.9	433	1 UKMS	u-plasminogen acti
7	241	47.3	431	2 JS0599	t-plasminogen acti
8	241	47.3	477	1 A34369	t-plasminogen acti
9	241	47.3	477	2 JS0598	t-plasminogen acti
10	226	44.3	291	2 JS0598	t-plasminogen acti
11	226	44.3	562	1 UKHUT	t-plasminogen acti
12	225	44.1	559	1 A35029	t-plasminogen acti
13	220	43.1	394	2 JS0600	t-plasminogen acti
14	219	42.9	559	1 A29941	t-plasminogen acti
15	213	41.8	477	2 JS0597	t-plasminogen acti
16	209.5	41.1	558	2 JC5878	plasma hyaluronan-
17	198.5	38.9	434	1 A35005	plasma hyaluronan-
18	193.5	37.9	560	1 JC4795	plasma hyaluronan-
19	193	37.8	655	1 A46688	hepatocyte growth
20	180.5	35.4	603	2 S28941	coagulation factor
21	170.5	33.4	615	1 KFHU12	coagulation factor
22	161	31.6	1420	2 A32869	apolipoprotein(a)
23	156	30.6	120	2 E61545	plasmin (EC 3.4.21
24	156	30.6	4548	1 S00657	apoptein(a) (EC
25	155.5	30.5	593	2 S45281	coagulation factor
26	153	30.0	89	2 A60140	plasmin (EC 3.4.21
27	150	29.4	123	2 G61545	plasmin (EC 3.4.21
28	149.5	29.3	169	2 A40522	plasmin (EC 3.4.21
29	149	29.2	812	1 PLBO	plasmin (EC 3.4.21

30	148.5	29.1	810	2	I46260	plasmin (EC 3.4.21
31	148	29.0	790	1	PLPG	plasmin (EC 3.4.21
32	147.5	28.9	460	2	B61545	plasmin (EC 3.4.21
33	147.5	28.9	943	2	B45082	neurotrophic recep
34	146	28.6	937	2	A45082	neurotrophic recep
35	144	28.2	810	1	PLHU	plasmin (EC 3.4.21
36	143	28.0	806	2	T18840	hypothetical prote
37	142.5	27.9	711	1	A47136	macrophage-stimula
38	141	27.6	810	2	B30848	plasmin (EC 3.4.21
39	139	27.3	685	1	A48289	neurotrophic recep
40	137.5	27.0	728	1	A60185	hepatocyte growth
41	135.5	26.6	728	1	A35644	hepatocyte growth
42	134	26.3	455	2	A61545	plasmin (EC 3.4.21
43	133.5	26.2	728	1	JH0579	hepatocyte growth
44	132.5	26.0	411	2	I51285	hepatocyte growth
45	132.5	26.0	2869	2	T18518	apolipoprotein(a)

ALIGNMENTS

RESULT 1

UKHU

u-plasminogen activator (EC 3.4.21.73) precursor [validated] - human
N;Alternate names: cellular plasminogen activator; urokinase; urokinase-type plasminogen a
N;Contains: urokinase-type plasminogen activator chain A; urokinase-type plasminogen a
in form

C;Species: Homo sapiens (man)
C;Date: 17-Dec-1982 #sequence revision 04-Dec-1986 #text change 15-Sep-2000
C;Accession: A00931; I52209; J70102; A37561; I38102; S65783; A37562; A37563; A37564; A
R;Riccio, A.; Grimaldi, G.; Verde, P.; Sebastio, G.; Boast, S.; Blasi, F.
Nucleic Acids Res. 13, 2759-2771, 1985
A;Title: The human urokinase-plasminogen activator gene and its promoter.
A;Reference number: A00931; MUID:85215647; PMID:2987867
A;Accession: A00931
A;Molecule type: DNA
A;Residues: 1-431 <RTC>
A;Cross-references: GB:X02419; NID:G37601; PIDN:CAA26268.1; PID:G1834524
A;Note: the authors translated the codon ATG for residue 214 as Ile
R;Nagamine, Y.; Pearson, D.; Grattan, M.
Biochem. Biophys. Res. Commun. 132, 563-569, 1985

A;Title: Exon-intron boundary sliding in the generation of two mRNAs coding for porcin
A;Reference number: I52209; MUID:86050639; PMID:3933505
A;Accession: I52209
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 145-161 <NAG1>
A;Cross-references: GB:X03027; NID:G340174; PIDN:AAA61257.1; PID:G340175
R;Nagai, M.; Hiramatsu, R.; Kaneda, T.; Hayasuke, N.; Arimura, H.; Nishida, M.; Suyama
Gene 36, 183-188, 1985
A;Title: Molecular cloning of cDNA coding for human preprourokinase.
A;Reference number: J70102; MUID:86056954; PMID:2415429
A;Accession: J70102
A;Molecule type: mRNA
A;Residues: 1-213, 'I', 215-431 <NAG2>
A;Cross-references: GB:X03226; NID:G340155; PIDN:AA097138.1; PID:G340158; GB:D00244; N
R;Verde, P.; Stoppelli, M.P.; Galeffi, P.; Di Nocera, P.; Blasi, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 4727-4731, 1984

A;Title: Identification and primary sequence of an unspliced human urokinase poly(A) +
A;Reference number: A37561; MUID:84272706; PMID:6589620
A;Accession: A37561
A;Molecule type: mRNA
A;Residues: 66-431 <VER>
A;Cross-references: GB:D00244; NID:G220138
R;Jacobs, P.; Cravador, A.; Loriau, R.; Brockly, F.; Colau, B.; Chuchana, P.; van Elise
DNA 4, 139-146, 1985

A;Title: Molecular cloning, sequencing, and expression in Escherichia coli of human pr
A;Reference number: I38102; MUID:85203359; PMID:3888571
A;Accession: I38102
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-150, 'W', 152-213, 'I', 215-385, 'C', 387-429, 'V', 431 <JAC>
A;Cross-references: EMBL:X02760; NID:G35297; PIDN:CAA26535.1; PID:G35298

R; Yoshimoto, M.; Ushiyama, Y.; Sakai, M.; Tamaki, S.; Hara, H.; Takahashi, K.; Sawasaki, K. *Biochim. Biophys. Acta* 1293, 88-89, 1996
 A; Title: Characterization of single chain urokinase-type plasminogen activator with a novel structure.
 A; Reference number: S65783; MUID:96186279; PMID:8652631
 A; Accession: S65783
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 21-140, 'L', 142-213, 'I', 215-431 <YOS>
 A; Cross-references: EMBL:D11143; NID:91311467; PIDN:BA001919.1; PID:g1199928
 R; Gunzler, W.A.; Steffens, G.J.; Otting, F.; Kim, S.M.A.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165, 1982
 A; Title: The primary structure of high molecular mass urokinase from human urine.
 A; Reference number: A37562; MUID:83055084; PMID:6754569
 A; Accession: A37562
 A; Molecule type: Protein
 A; Residues: 21-177 <GUN>
 R; Schaller, J.; Nick, H.; Rickli, E.E.; Gillesen, D.; Lergier, W.; Studer, R.O.
 Eur. J. Biochem. 125, 251-257, 1982
 A; Title: Human low-molecular-weight urinary urokinase. Partial characterization and preliminary structure.
 A; Reference number: A37563; MUID:83003608; PMID:6749491
 A; Accession: A37563
 A; Molecule type: Protein
 A; Residues: 156-176; 179-193, 'T', 195, 'T', 197-224 <SCH>
 R; Steffens, G.J.; Gunzler, W.A.; Otting, F.; Frankus, E.; Flohe, L.
 Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058, 1982
 A; Title: The complete amino acid sequence of low molecular mass urokinase from human urine.
 A; Reference number: A37564; MUID:83055099; PMID:6754572
 A; Accession: A37564
 A; Molecule type: Protein
 A; Residues: 158-410 <STE>
 R; Kentzer, E.J.; Buko, A.; Menon, G.; Sarin, V.K.
 Biochem. Biophys. Res. Commun. 171, 401-406, 1990
 A; Title: Carbohydrate composition and presence of a fucose-protein linkage in recombinant urokinase.
 A; Reference number: A35689; MUID:90365737; PMID:2393398
 A; Accession: A35689
 A; Molecule type: Protein
 A; Residues: 21-30, 'X', 32, 'X', 34-38, 'X', 40-43 <KEN>
 A; Note: Identification of a fucose and attempt to determine its attachment site.
 R; Rabbani, S.A.; Desjardins, J.; Bell, A.W.; Banville, D.; Mazar, A.; Henkin, J.; Goltzman, I.
 Biochem. Biophys. Res. Commun. 173, 1058-1064, 1990
 A; Title: An amino-terminal fragment of urokinase isolated from a prostate cancer cell line.
 A; Reference number: A36697; MUID:91097529; PMID:2125213
 A; Accession: A36697
 A; Molecule type: Protein
 A; Residues: 21-34 <RAB>
 R; Li, X.; Bokman, A.M.; Llinas, M.; Smith, R.A.G.; Dobson, C.M.
 Submitted to the Brookhaven Protein Data Bank, July 1993
 A; Reference number: A51255; PDB:1KDU
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Li, X.; Smith, R.A.G.; Dobson, C.M.
 Biochemistry 31, 9562-9571, 1992
 A; Title: Sequential (1)H NMR assignments and secondary structure of the kringle domain of urokinase.
 A; Reference number: A44375; MUID:93003110; PMID:1327118
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
 R; Hansen, A.P.; Petros, A.M.; Meadows, R.P.; Nettesheim, D.G.; Mazar, A.P.; Olejniczak, Z.
 Submitted to the Brookhaven Protein Data Bank, January 1994
 A; Reference number: A66822; PDB:1URK
 A; Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
 R; Spraggon, G.S.; Phillips, C.; Nowak, U.K.; Ponting, C.P.; Saunders, D.; Dobson, C.M.;
 submitted to the Brookhaven Protein Data Bank, July 1995
 A; Reference number: A66058; PDB:1LMW
 A; Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 168-175; 179-426
 C; Comment: This enzyme is found in urine in a high molecular mass form, consisting of A and B chains.
 C; Comment: Urokinase-type plasminogen activator proteolytically activates plasminogen, a zymogen of plasmin.
 C; Genetics:
 A; Gene: GDB:PLAU
 A; Cross-references: GDB:119497; OMIM:191840
 A; Map position: 10q24-10q24
 A; Introns: 19/3; 29/1; 65/1; 123/2; 154/1; 227/2; 277/1; 324/1; 373/3
 C; Function:
 A; Description: proteolytically activates plasminogen
 A; Pathway: fibrinolysis
 C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try

C; Keywords: fibrinolysis; glycoprotein; heterodimer; hydrolase; kringle; serine protease
 F; 1-20/Domain: signal sequence #status predicted <SIG>
 F; 21-431/Product: urokinase-type plasminogen activator, single chain form #status predicted <SIG>
 F; 21-177/Product: urokinase-type plasminogen activator chain A #status experimental <M>
 F; 31-62/Domain: EGF homology <EGF>
 F; 70-151/Domain: kringle homology <KRG>
 F; 156-177/Product: urokinase-type plasminogen activator chain A1 #status experimental
 F; 179-431/Product: urokinase-type plasminogen activator chain B #status experimental <KRG>
 F; 179-419/Domain: trypsin homology <TRY>
 F; 31-39, 33-51, 53-62, 70-151, 91-133, 122-146, 168-299, 209-225, 217-288, 313-382, 345-361, 372-382/Binding site: carbohydrate (Thr) (covalent) #status predicted
 F; 38/Binding site: carbohydrate (Thr) (covalent) #status experimental
 F; 178-179/Cleavage site: Lys-Ile (plasma) #status experimental
 F; 224, 275, 376/Active site: His, Asp, Ser #status experimental
 F; 322/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 100.0%; Score 510; DB 1; Length 431;

Best Local Similarity 100.0%; Pred. No. 4.3e-48;

Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLTQTYHAHRSALQLGLGKHNYCRNPDN 60

Db 68 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLTQTYHAHRSALQLGLGKHNYCRNPDN 127

Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88

Db 128 RRRPWCYVQGLKPLVQECMVHDCADGK 155

RESULT 2

UKBAY

u-plasminogen activator (EC 3.4.21.73) precursor - yellow baboon

C; Species: Papio cynocephalus, Papio hamadryas cynocephalus (yellow baboon)

C; Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999

C; Accession: S14687; S08651

R; Au, Y.P.T.; Wang, T.W.; Clowes, A.W.

Nucleic Acids Res. 18, 3411, 1990

A; Title: Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator

A; Reference number: S14687; MUID:90287734; PMID:2113276

A; Accession: S14687

A; Molecule type: mRNA

A; Residues: 1-433 <AU>

A; Cross-references: EMBL:X51935; NID:G38130; PIDN:CAA36200.1; PID:G38131

C; Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t

C; Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase

F; 1-20/Domain: signal sequence #status predicted <SIG>

F; 21-176/Product: plasminogen activator chain A #status predicted <ACH>

F; 30-61/Domain: EGF homology <EGF>

F; 69-150/Domain: kringle homology <KRG>

F; 178-433/Product: plasminogen activator chain B #status predicted <BCH>

F; 178-421/Domain: trypsin homology <TRY>

F; 167-238, 208-224, 216-287, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted

F; 223, 274, 378/Active site: His, Asp, Ser #status predicted

F; 324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 93.1%; Score 475; DB 1; Length 433;

Best Local Similarity 94.3%; Pred. No. 2.9e-44;

Matches 83; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLTQTYHAHRSALQLGLGKHNYCRNPDN 60

Db 67 KTCYEGNGHFYRGKASTDTMGRCPLPWNATVLTQTYHAHRSALQLGLGKHNYCRNPDN 126

Qy 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88

Db 127 RRRPWCYVQGLKPLVQECMVHDCADGK 154

RESULT 3

UKPG

u-plasminogen activator (EC 3.4.21.73) precursor - pig

N; Alternate names: uPA

C; Species: Sus scrofa domestica (domestic pig)

C; Date: 04-Dec-1986 #sequence_revision 17-Mar-1987 #text_change 07-Aug-1998

C/Accession: A00932
R/Nagamine, Y.; Pearson, D.; Altus, M.S.; Reich, B.
Nucleic Acids Res. 12, 9525-9541, 1984
A/Title: cDNA and gene nucleotide sequence of porcine plasminogen activator.
A/Reference number: A00932; MUID:85087954; PMID:6096832
A/Accession: A00932
A/Molecule type: DNA
A/Residues: 1-240, 'H', 242-442 <NAG1>
A/Experimental source: kidney cell line LLC-PK1
R/Nagamine, Y.
Submitted to the Protein Sequence Database, December 1986
A/Reference number: A37566
A/Contents: annotation; correction to residue 241
C/Genetics:
A/Introns: 19/3; 31/1; 67/1; 125/2; 165/1; 238/2; 288/1; 335/1; 384/3
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-198/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F/33-64/Domain: EGF homology <EGF>
F/72-153/Domain: kringle homology <KRG>
F/190-442/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F/190-430/Domain: trypsin homology <TRY>
F/152/Binding site: Carboxylate (Asn) (covalent) #status predicted
F/179-310, 220-236, 228-299, 324-393, 356-372, 383-411/Disulfide bonds: #status predicted
F/235, 286, 387/Active site: His, Asp, Ser #status predicted

Query Match 82.4%; Score 420; DB 1; Length 442;
Best Local Similarity 81.8%; Pred. No. 3.1e-38;
Matches 72; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 KTCYEGNGHGYRCKASTDTMGRPCLPWNSATVLQQTVAHRSALQGLGKHNYCRNPDN 60
Db 70 QTCFEGNGHGYRCKANTDGRPCLPWNSATVLLNTHAHPDALQGLGKHNYCRNPDN 129

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 130 QRRPWCYVQVGLKPLVQECMVPCSGGE 157

RESULT 4
S18932
u-plasminogen activator (EC 3.4.21.73) precursor - rat
N/Alternate names: plasminogen activator, urokinase-type; urinary plasminogen activator
C/Species: Rattus norvegicus (Norway rat)
C/Date: 18-Oct-1989 #sequence_revision 10-Feb-1995 #text_change 18-Jun-1999
C/Accession: S24604; I60186; I53472; S18932
R/Rabbani, S.A.
submitted to the EMBL Data Library, April 1992
A/Reference number: S24604
A/Accession: S24604
A/Molecule type: mRNA
A/Residues: 1-15, 'H', 17-23, 'G', 25-331, 'N', 333-432 <RAB>
A/Cross-references: EMBL:X65651; NID:G57456; PIDN:CAA46601.1; PID:G57457
A/Experimental source: tissue kidney
R/Henderson, B.R.; Tansey, W.P.; Phillips, S.M.; Ramshaw, I.A.; Kefford, R.F.
Cancer Res. 52, 2489-2496, 1992
A/Title: Transcriptional and posttranscriptional activation of urokinase plasminogen act
A/Reference number: I60186; MUID:92233409; PMID:1568219
A/Accession: I60186
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-432 <RES>
A/Cross-references: EMBL:X63434; NID:G57465; PIDN:CAA45028.1; PID:G57466
A/Experimental source: strain Fischer 344; tissue mammary
R/Ragno, P.; Cassano, S.; Begen, J.; Kessler, C.; Blasi, F.; Rossi, G.
FEBS Lett. 306, 193-198, 1992
A/Title: The receptor for the plasminogen activator of urokinase type is up-regulated in
A/Reference number: I53472; MUID:92339549; PMID:1321734
A/Accession: I53472
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 31-62 <RE2>

A/Cross-references: EMBL:X66907; NID:G396200; PIDN:CAA47356.1; PID:G938279
C/Genetics:
A/Gene: uPA
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F/1-19/Domain: signal sequence #status predicted <SIG>
F/20-177/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
F/31-62/Domain: EGF homology <EGF>
F/70-151/Domain: kringle homology <KRG>
F/179-432/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
F/179-420/Domain: trypsin homology <TRY>
F/168-300, 210-226, 218-289, 314-383, 346-362, 373-401/Disulfide bonds: #status predicted
F/225, 276, 377/Active site: His, Asp, Ser #status predicted

Query Match 79.6%; Score 406; DB 1; Length 432;
Best Local Similarity 79.5%; Pred. No. 1e-36;
Matches 70; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHGYRCKASTDTMGRPCLPWNSATVLQQTVAHRSALQGLGKHNYCRNPDN 60
Db 68 KTCYHGNGHGYRCKANTDGRPCLPWNSATVLLNTHAHPDALQGLGKHNYCRNPDN 127

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 128 QRRPWCYVQVGLKPLVQECMVQDCSLK 155

RESULT 5
JN0560
u-plasminogen activator (EC 3.4.21.73) precursor - bovine
N/Alternate names: uPA
C/Species: Bos primigenius taurus (cattle)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C/Accession: JN0560
R/Kraetzschmar, J.; Haendler, B.; Kojima, S.; Rifkin, D.B.; Schleuning, W.D.
Gene 125, 177-183, 1993
A/Title: Bovine urokinase-type plasminogen activator and its receptor: cloning and ind
A/Reference number: JN0560; MUID:93216119; PMID:8385052
A/Accession: JN0560
A/Molecule type: mRNA
A/Residues: 1-433 <KRA>
A/Cross-references: GB:L03546; NID:G163800; PIDN:AAA51419.1; PID:G163801
C/Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; t
C/Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
F/1-20/Domain: signal sequence #status predicted <SIG>
F/21-179/Product: plasminogen activator chain A #status predicted <MA1>
F/33-64/Domain: EGF homology <EGF>
F/72-153/Domain: kringle homology <KRG>
F/181-433/Product: plasminogen activator chain B #status predicted <MA2>
F/181-421/Domain: trypsin homology <TRY>
F/170-301, 211-227, 219-290, 315-384, 347-363, 374-402/Disulfide bonds: #status predicted
F/226, 277, 378/Active site: His, Asp, Ser #status predicted

Query Match 77.6%; Score 396; DB 1; Length 433;
Best Local Similarity 76.1%; Pred. No. 1.3e-35;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHGYRCKASTDTMGRPCLPWNSATVLQQTVAHRSALQGLGKHNYCRNPDN 60
Db 70 KTCYQNGHGYRCKANRDLGRPCLPWNSATVLLNTHAHPDALQGLGKHNYCRNPDN 129

QY 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 130 QRRPWCYVQVGLKPLVQECMVQDCSVGK 157

RESULT 6
UKWS
u-plasminogen activator (EC 3.4.21.73) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C/Accession: A29420; A24615

R;Degen, S.J.F.; Heckel, J.L.; Reich, E.; Degen, J.L.
 Biochemistry 26, 8270-8279, 1987
 A;Title: The murine urokinase-type plasminogen activator gene.
 A;Reference number: A29420; MUID:88163489; PMID:2831940
 A;Accession: A29420
 A;Molecule type: DNA
 A;Residues: 1-433 <DEG>
 A;Cross-references: GB:M17922; NID:g202296; PIDN:AAA40539.1; PID:g202297
 R;Belin, D.; Vassalli, J.D.; Combepine, C.; Godeau, F.; Nagamine, Y.; Reich, E.; Kocher, Eur. J. Biochem. 148, 225-232, 1985
 A;Title: Cloning, nucleotide sequencing and expression of cDNAs encoding mouse urokinase
 A;Reference number: A24615; MUID:85179474; PMID:2985383
 A;Accession: A24615
 A;Molecule type: mRNA
 A;Residues: 1-433 <BEL>
 A;Cross-references: GB:X02389; NID:g55127; PIDN:CAA26231.1; PID:g55128
 C;Genetics:
 A;Introns: 19/3; 30/1; 66/1; 124/2; 155/1; 229/2; 279/1; 326/1; 375/3
 C;Superfamily: urokinase-type plasminogen activator; EGF homology; kringle homology; try
 C;Keywords: glycoprotein; heterodimer; hydrolase; kringle; serine proteinase
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-178/Product: urokinase-type plasminogen activator chain A #status predicted <ACH>
 F;32-63/Domain: EGF homology <EGF>
 F;71-152/Domain: kringle homology <KRG>
 F;180-433/Product: urokinase-type plasminogen activator chain B #status predicted <BCH>
 F;180-421/Domain: trypsin homology <TRY>
 F;169-301,211-227,219-290,315-384,347-363,374-402/disulfide bonds: #status predicted
 F;226,277,378/Active site: His, Asp, Ser #status predicted
 Query Match 75.9%; Score 387; DB 1; Length 433;
 Best Local Similarity 72.7%; Pred. No. 1.2e-34;
 Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 Qy 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTVLOQTYHAHRSDALQGLGKHNYCRNPDN 60
 Db 69 KTCYHGNGDSYRGKASTDTMGRCPLPWNSTVLOQTYHAHRSDALQGLGKHNYCRNPDN 128
 Qy 61 RRPWCYVQVGLKPLVQECMVHDCADGK 88
 Db 129 QKRPWCYVQIGLRFQVQECMVHDCSLK 156
 RESULT 7
 JS0599
 t-plasminogen activator (EC 3.4.21.68) beta precursor - common vampire bat
 N;Alternate names: tissue plasminogen activator
 C;Species: Desmodus rotundus (common vampire bat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: JS0599
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; Don
 Gene 105, 229-237, 1991
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
 A;Reference number: JS0597; MUID:92039036; PMID:1937019
 A;Accession: JS0599
 A;Molecule type: mRNA
 A;Residues: 1-431 <KRA>
 A;Cross-references: GB:M63989; NID:g166076; PIDN:AAA31594.1; PID:g166077
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-431/Product: plasminogen activator beta #status predicted <PLA>
 F;41-74/Domain: EGF homology <EGF>
 F;82-163/Domain: kringle homology <KRG>
 F;180-425/Domain: trypsin homology <TRY>
 F;41-52,46-63,65-74,82-163,103-145,134-158,168-299,211-227,219-288,313-388/Disulfide bon
 F;135,352/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;179-180/Cleavage site: His-Ser (plasmin) #status predicted
 F;226,275,382/Active site: His, Asp, Ser #status predicted
 F;345-361,378-406/Disulfide bonds: #status predicted
 Query Match 47.3%; Score 241; DB 2; Length 431;
 Best Local Similarity 50.0%; Pred. No. 1.1e-18;

Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 Qy 2 TCYEGNGHFYRGKASTDTMGRCPLPWNSTVLOQTYHAHRSDALQGLGKHNYCRNPDN 61
 Db 81 TCYKQDQGVYRGTWSTESGAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDN 140
 Qy 62 RRPWCYVQVGLKPLVQECMVHDC 85
 Db 141 SKPCYVIKASKFILEFCSPVPCS 164
 RESULT 8
 A34369
 t-plasminogen activator (EC 3.4.21.68) precursor - false vampire bat (Megaderma lyra)
 C;Species: Megaderma lyra
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C;Accession: A34369
 R;Gardell, S.J.; Duong, L.T.; Diehl, R.E.; York, J.D.; Hare, T.R.; Register, R.B.; Jac
 J. Biol. Chem. 264, 17947-17952, 1989
 A;Title: Isolation, characterization, and cDNA cloning of a vampire bat salivary plasm
 A;Reference number: A34369; MUID:90036867; PMID:2509450
 A;Accession: A34369
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-477 <GAR>
 A;Cross-references: GB:J05082; NID:g166080; PIDN:AAA31596.1; PID:g166081
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <1FA>
 F;87-120/Domain: EGF homology <EGF>
 F;128-209/Domain: kringle homology <KRG>
 F;226-471/Domain: trypsin homology <TRY>
 F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359
 F;272,321,428/Active site: His, Asp, Ser #status predicted
 Query Match 47.3%; Score 241; DB 1; Length 477;
 Best Local Similarity 50.0%; Pred. No. 1.2e-18;
 Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;
 Qy 2 TCYEGNGHFYRGKASTDTMGRCPLPWNSTVLOQTYHAHRSDALQGLGKHNYCRNPDN 61
 Db 127 TCYKQDQGVYRGTWSTESGAQCINWNSNLLTRTYNGRRSDAITLGLGNHNYCRNPDN 186
 Qy 62 RRPWCYVQVGLKPLVQECMVHDC 85
 Db 187 SKPCYVIKASKFILEFCSPVPCS 210
 RESULT 9
 JS0598
 t-plasminogen activator (EC 3.4.21.68) alpha-2 precursor - common vampire bat
 N;Alternate names: tissue plasminogen activator
 C;Species: Desmodus rotundus (common vampire bat)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
 C;Accession: JS0598
 R;Kraetzschmar, J.; Haendler, B.; Langer, G.; Boidol, W.; Bringmann, P.; Alagon, A.; I
 Gene 105, 229-237, 1991
 A;Title: The plasminogen activator family from the salivary gland of the vampire bat I
 A;Reference number: JS0597; MUID:92039036; PMID:1937019
 A;Accession: JS0598
 A;Molecule type: mRNA
 A;Residues: 1-477 <KRA>
 A;Cross-references: GB:M63988; NID:g166074; PIDN:AAA31593.1; PID:g166075
 C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
 C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-36/Domain: propeptide #status predicted <PRO>
 F;37-477/Product: plasminogen activator alpha-2 #status predicted <PLA>
 F;42-79/Domain: fibronectin type I repeat homology <1FA>
 F;87-120/Domain: EGF homology <EGF>

```
F;128-209/Domain: kringle homology <KR>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,78-97,98-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359-444
F;185,398/Binding site: carboxylate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His, Asp, Ser #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match          47.3%; Score 241; DB 2; Length 477;
Best Local Similarity 50.0%; Pred. No. 1,2e+18;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

QY      2 TCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTTHAHRSDALQLGLGKHNYCRNPDDNR 61
Db      127 TCYKDQGVTYRGTWSTSESAQAICNNNLLTERTYNGRRSDAITLGLGNHNYCRNPDDN 186
               :|||||:::|:|||||
QY      62 RRPWCYYQVGLKPLVQECMVHDCA 85
Db      187 SKPCWYVIKASKFILEFCSPVCVS 210
               :|||||:::|:|||||

RESULT 10
I38098
t-plasminogen activator precursor, inactive endothelial splice form - human
N;Alternate names: tissue plasminogen activator
C;Species: Homo sapiens (man)
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 22-Jun-1999
C;Accession: I38098; S01678
R;Siebert, P.D.; Fong, K.
Nucleic Acids Res. 15, 1086, 1990
A;Title: Variant tissue type plasminogen activator (PLAT) cDNA obtained from human endothelial cells
A;Reference number: I38098; UID:90192129; PMID:1969145
A;Accession: I38098
A;Status: translated from GB/EMBL/DDBB
A;Molecule type: mRNA
A;Residues: 1-291 <SIE>
A;Cross-references: EMBL:X13097; NID:g35282; PIDN:CAA31489.1; PID:g35283
C;Comment: For the main splice form, see PIR:UKHUT. This form probably does not have pro
C;Genetics:
A;Gene: GDB:PLAT
A;Cross-references: GDB:I19496; OMIM:173370
A;Map position: 9p12-9p12
A;Intons: 24/3; 39/1; 85/1; 122/1; 180/2; 211/1; 268/2
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: alternative splicing; fibrinolysis; glycoprotein; kringle
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-32/Domain: propeptide #status predicted <PRO>
F;33-291/Product: t-plasminogen activator, inactive endothelial splice form #status pred
F;41-78/Domain: fibronectin type I repeat homology <1FA>
F;86-119/Domain: EGF homology <EGF>
F;127-208/Domain: kringle homology <KR1>
F;215-291/Domain: kringle homology #status atypical <KR2>
F;41-71,69-78,86-97,91-108,110-119,127-208,148-190,179-203/disulfide bonds: #status pred

Query Match          44.3%; Score 226; DB 2; Length 291;
Best Local Similarity 47.7%; Pred. No. 3.3e+17;
Matches 41; Conservative 8; Mismatches 37; Indels 0; Gaps 0;

QY      2 TCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTTHAHRSDALQLGLGKHNYCRNPDDNR 61
Db      126 TCYEDQGISYRGTWSTAESAECTNNSSALAQNAYSGRRPDRAIRLGLGNHNYCRNPDRD 185
               |||||   |||||   |||||   |||||   |||||   |||||   |||||   |||||
QY      62 RRPWCYYQVGLKPLVQECMVHDCA 87
Db      186 SKPCWYVFKAGYSBFCSTPACSEG 211
               :|||||:::|:|||||

RESULT 11
UKHUT
t-plasminogen activator (EC 3.4.21.68) precursor [validated] - human
N;Alternate names: t-PA; tissue plasminogen activator
C;Species: Homo sapiens (man)
C;Date: 14-Nov-1983 #sequence_revision 14-Nov-1993 #text_change 08-Dec-2000
C;Accession: A94004; A35229; J070562; A93293; S02125; A91343; A93951; A91322; A54645; I60
```

R;Ny, T.; Elgh, F.; Lund, B.
Proc. Natl. Acad. Sci. U.S.A. 81, 5355-5359, 1984
A:Title: The structure of the human tissue-type plasminogen activator gene: correlation
A:Reference number: A94004; MUID:84298137; PMID:6089198
A:Accession: A94004
A:Molecule type: DNA
A:Residues: 1-562 <NVT>
A:Cross-references: GB:L00141
A:Note: the codon given for residue 93 (ACC) is inconsistent with the authors' translation
R;Frieznier Degen, S.J.; Rajput, B.; Reich, E.
J. Biol. Chem. 261, 6972-6985, 1986
A:Title: The human tissue plasminogen activator gene.
A:Reference number: A23529; MUID:86196143; PMID:3009482
A:Accession: A23529
A:Molecule type: DNA
A:Residues: 1-562 <DEG>
A:Cross-references: GB:K03021; NID:g339817; PIDN:AAA98809.1; PID:g339818
R;Itagaki, Y.; Yasuda, H.; Morinaga, T.; Mitsuda, S.; Higashio, K.
Agric. Biol. Chem. 55, 1225-1232, 1991
A:Title: Purification and characterization of tissue plasminogen activator secreted by
A:Reference number: JT0562; MUID:91291340; PMID:1368681
A:Accession: JT0562
A:Molecule type: mRNA
A:Residues: 31-562 <ITA>
A:Cross-references: DDBJ:D01096; NID:G220128; PIDN:BAA00881.1; PID:g441174
A:Experimental source: embryonic lung fibroblast IMR-90 cells
A:Note: part of this sequence, including the amino end of the mature protein, was confirmed
R;Pennica, D.; Holmes, W.E.; Kohr, W.J.; Harkins, R.N.; Vehar, G.A.; Ward, C.A.; Benne
Nature 301, 214-221, 1983
A:Title: Cloning and expression of human tissue-type plasminogen activator cDNA in Escherichia
A:Reference number: A93293; MUID:83115262; PMID:6337343
A:Accession: A93293
A:Molecule type: mRNA
A:Residues: 1-562 <PEN>
A:Cross-references: GB:L00141
A:Experimental source: melanoma cells
R;Sasaki, H.; Saito, Y.; Hayashi, M.; Otsuka, K.; Niwa, M.
Nucleic Acids Res. 16, 5695, 1988
A:Title: Nucleotide sequence of the tissue-type plasminogen activator cDNA from human
A:Reference number: S02125; MUID:88262579; PMID:3133640
A:Accession: S02125
A>Status: translation not shown
A:Molecule type: mRNA
A:Residues: 1-562 <SAS>
A:Cross-references: EMBL:X07393; NID:g37243; PIDN:CAA30302.1; PID:g37244
R;Kagitani, H.; Tagawa, M.; Hatanaka, K.; Ikari, T.; Saito, A.; Bando, H.; Okada, K.; I
FEBS Lett. 189, 145-149, 1985
A:Title: Expression in Escherichia coli of finger-domain lacking tissue-type plasminogen
A:Reference number: A91343; MUID:85285620; PMID:3896853
A:Accession: A91343
A:Molecule type: mRNA
A:Residues: 1-38,'G','86-433','E',435-562 <KAG>
R;Edlund, T.; Ny, T.; Ranby, M.; Heden, L.O.; Palm, G.; Holmgren, E.; Josephson, S.
Proc. Natl. Acad. Sci. U.S.A. 80, 349-352, 1983
A:Title: Isolation of cDNA sequences coding for a part of human tissue plasminogen act
A:Reference number: A93951; MUID:83169656; PMID:6572897
A:Accession: A93951
A:Molecule type: mRNA
A:Residues: 251-358 <EDL>
R;Pohl, G.; Kallstrom, M.; Bergsdorf, N.; Wallen, P.; Jornvall, H.
Biochemistry 23, 3701-3707, 1984
A:Title: Tissue plasminogen activator: peptide analyses confirm an indirectly derived
A:Reference number: A90488; MUID:85000468; PMID:6433976
A:Contents: annotation; melanoma cells, partial sequence of residues 36-562, active an
R;Pohl, G.; Kaplan, L.; Einarsson, M.; Wallen, P.; Jornvall, H.
FEBS Lett. 168, 29-32, 1984
A:Title: Differences between uterine and melanoma forms of tissue plasminogen activator:
A:Reference number: A91322; MUID:84158956; PMID:6538514
A:Accession: A91322

Query Match 44.3%; Score 226; DB 1; Length 562;

C;Accession: JS0600
R;Kraetzschmar, J.: Haendler, B.: Boidol, W.: Bringmann, P.: Alagon, A.: E

Gene 105, 229-237, 1991
A;Title: The plasminogen activator family from the salivary gland of the vampire bat Des
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0600
A;Molecule type: mRNA
A;Residues: 1-394 <KRA>
A;Cross-references: GB:M63990; NID:gl66078; PIDN:AAA1595.1; PID:gl66079
A;Note: the authors translated the codon A1C for residue 75 as Thr
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-394/Product: plasminogen activator gamma #status predicted <PLA>
F;45-126/Domain: kringle homology <KR>
F;143-388/Domain: trypsin homology <TRY>
F;45-126,66-108,97-121,131-262,174-190,182-251,276-351,308-324,341-369/Disulfide bonds:
F;142-143/Cleavage site: His-Ser (plasmin) #status predicted
F;189,238,345/Active site: His, Asp, Ser #status predicted
F;315/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 43.1%; Score 220; DB 2; Length 394;
Best Local Similarity 45.2%; Pred. No. 2e-16;
Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGRKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDR 61
Db 44 TCYKDGVTYRGWTSSTSGAQCINNNSLLIRTYNGRMPKPAVLGLGNHNYCRNPDR 103

QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 104 SKPWCYVIRAKRTSESVCSPVCS 127

RESULT 14
A29941
t-plasminogen activator (EC 3.4.21.68) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A29941; S48205; S48207; S48206
J.Rickles, R.J.; Darrow, A.L.; Strickland, S.
J. Biol. Chem. 263, 1563-1569, 1988
A;Title: Molecular cloning of complementary DNA to mouse tissue plasminogen activator mR
A;Reference number: A29941; MUID:88087303; PMID:2826484
A;Accession: A29941
A;Molecule type: mRNA
A;Residues: 1-559 <RIC>
A;Cross-references: GB:J03520; NID:G202109; PIDN:AAA0470.1; PID:G202110
R.Lijnen, H.R.; van Hoef, B.; Beelen, V.; Collen, D.
Eur. J. Biochem. 224, 863-871, 1994
A;Title: Characterization of the murine plasma fibrinolytic system.
A;Reference number: S48202; MUID:95010076; PMID:7523120
A;Accession: S48205
A;Molecule type: protein
A;Residues: 33-37, 'X', 39-40 <LIJ>
A;Accession: S48207
A;Molecule type: protein
A;Residues: 309-316 <LI2>
A;Accession: S48206
A;Molecule type: protein
A;Residues: 33-37, 'X', 39-40 <LIW>
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat hom
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-29/Domain: propeptide #status predicted <PRO>
F;30-559/Product: t-plasminogen activator #status predicted <MAT>
F;30-308/Product: t-plasminogen activator chain A #status predicted <ACH>
F;38-75/Domain: fibronectin type I repeat homology <1F1>
F;83-116/Domain: EGF homology <EGF>
F;124-205/Domain: kringle homology <KR1>
F;213-294/Domain: kringle homology <KR2>
F;309-559/Product: t-plasminogen activator chain B #status predicted <BCH>
F;309-553/Domain: trypsin homology <TRY>
F;38-68,66-75,83-94,88-105,107-116,124-145,145-187,176-200,213-294,234-276,265-289,297-4

F;149,481/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;308-309/Cleavage site: Arg-Ile (plasmin, trypsin) #status predicted
F;355,404,510/Active site: His, Asp, Ser #status predicted

Query Match 42.9%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred. No. 3.6e-16;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGRKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDR 61
Db 123 TCPEQGITYRGWTSSTSGAQCINNNSVLSLKPNARRPNAIKLGLGNHNYCRNPDR 182

QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 183 LKPWCYVFRAGKTYTFECPACPKGK 209

RESULT 15
JS0597
t-plasminogen activator (EC 3.4.21.68) alpha-1 precursor - common vampire bat
N;Alternate names: tissue plasminogen activator
C;Species: Desmodus rotundus (common vampire bat)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C;Accession: JS0597
R;Raetzschmar, J.; Haendler, B.; Langer, G.; Boigol, W.; Bringmann, P.; Alagon, A.; D
Gene 105, 229-237, 1991
A;Title: the plasminogen activator family from the salivary gland of the vampire bat D
A;Reference number: JS0597; MUID:92039036; PMID:1937019
A;Accession: JS0597
A;Molecule type: mRNA
A;Residues: 1-477 <KRA>
A;Cross-references: GB:M63987; NID:gl66070; PIDN:AAA1591.1; PID:gl66071
C;Superfamily: tissue plasminogen activator; EGF homology; fibronectin type I repeat h
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-36/Domain: propeptide #status predicted <PRO>
F;37-477/Product: plasminogen activator alpha-1 #status predicted <PLA>
F;42-79/Domain: fibronectin type I repeat homology <1FA>
F;87-120/Domain: EGF homology <EGF>
F;128-209/Domain: kringle homology <KR>
F;226-471/Domain: trypsin homology <TRY>
F;42-72,70-79,87-98,92-109,111-120,128-209,149-191,180-204,214-345,257-273,265-334,359
F;153,398/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;225-226/Cleavage site: His-Ser (plasmin) #status predicted
F;272,321,428/Active site: His, Asp, Ser #status predicted

Query Match 41.8%; Score 213; DB 2; Length 477;
Best Local Similarity 46.4%; Pred. No. 1.4e-15;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGRKASTDTWGRPCLPWNSATVLOQTYHAHRSALQGLGKHNYCRNPDR 61
Db 127 TCYEGQGVTYRGWTSSTSGAQCINNNSLLIRTYNGRMPDAFNLGLGNHNYCRNPNGA 186

QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 KPWCYVIRAKRTSESVCSPVCS 210

RESULT 16
JCS878
plasma hyaluronan-binding protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Mar-1998 #sequence_revision 11-Mar-1998 #text_change 16-Jul-1999
C;Accession: JCS878
R;Hashimoto, K.; Tobe, T.; Sumiya, J.; Saguchi, K.; Sano, Y.; Nakano, Y.; Choi-Miura, I
Biol. Pharm. Bull. 20, 1127-1130, 1997
A;Title: Cloning of the cDNA for a mouse homologue of human PHBP: A novel hyaluronan-b
A;Reference number: JCS878; MUID:98065239; PMID:9401717
A;Accession: JCS878
A;Molecule type: mRNA
A;Residues: 1-558 <HAS>
C;Comment: This protein acts as serine protease.

A:Title: Purification and characterization of a novel hyaluronan-binding protein (PHBP) r activator.
A:Reference number: JC4795; MUID:96425001; PMID:8827452
A:Accession: JC4795
A:Molecule type: mRNA
A:Residues: 1-560 <CHO>
A:Cross-references: GB:S83182; NID:G1836158; PIDN:AAB46909.1; PID:G1836159
A:Experimental source: plasma
A:Note: parts of this sequence, including the amino ends of the mature chains, were det
C:Genetics:
A:Gene: GDB:HABP2; HABP; PHBP; HGFAL
A:Cross-references: GDB:4573962
C:Complex: a disulfide-bonded heterodimer of chains produced from the same precursor; t
C:Superfamily: plasma hyaluronan-binding protein; EGF homology; kringle homology; trypsin
C:Keywords: chondroitin sulfate proteoglycan; glycoprotein; hyaluronic acid; hydrolase;
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-313/Product: plasma hyaluronan-binding protein, 50K chain #status predicted <50K>
F:77-108/Domain: EGF homology <EG1>
F:115-147/Domain: EGF homology <EG2>
F:154-187/Domain: EGF homology <EG3>
F:194-276/Domain: kringle homology <KRI>
F:334-550/Domain: trypsin homology <TRY>
F:343-516/Product: plasma hyaluronan-binding protein, catalytic chain #status predicted
F:54-207/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:77-88,182-97,99-108,115-125,120-136,138-147,154-165,159-176,178-187,194-276,215-257,24
F:362,405,509/Active site: His, Asp, Ser #status predicted

Query Match 37.98; Score 193.5; DB 1; Length 560;
Best Local Similarity 41.79; Pred. No. 2.2e-13;
Matches 35; Conservative 15; Mismatches 33; Indels 1; Gaps 1;

Qy 3 CYEGNGHFYRGKASTDTMGRPCLEPWNSALVLOQTYHAHRSDALQLGLGKHNYCRNPDDR 62
Db 194 CYVGDGYSYRGKMRTVQNHACLYWNSHLILLQENYMFEMEDAETHGIGEHNFCDPDAE 253

Qy 63 RPWCYQVGLKPLVQB-CMVHDC A 85
Db 254 KPWCFIKVTNDKVKWEYCDVSACS 277

RESULT 19
A46688
hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human
C:Species: Homo sapiens (man)
C:Date: 21-Sep-1993 #sequence revision 25-Aug-1995 #text_change 08-Dec-2000
C:Accession: A46688
R:Miyaazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.
J. Biol. Chem. 268, 10024-10028, 1993
A:Title: Molecular cloning and sequence analysis of the cDNA for a human serine proteas
d coagulation factor XII.
A:Reference number: A46688; MUID:93252878; PMID:7683665
A:Accession: A46688
A:Molecule type: mRNA
A:Residues: 1-655 <MY>
A:Cross-references: DDBJ:D14012; NID:G219680; PIDN:BAAO3113.1; PID:G219681
A:Experimental source: liver (mRNA); serum (protein)
A:Note: sequence extracted from NCBI backbone (NCBI:N131227, NCBI:P131228)
A:Note: parts of the sequence, including the amino ends of the heavy and light chains,
C:Genetics:
A:Gene: GDB:HGFAC; HGFPA; HGFAP
A:Cross-references: GDB:9954514
A:Map position: 4p16-4p16
C:Function:
A:Description: activates hepatocyte growth factor by specific proteolytic cleavage
A:Pathway: tissue repair and regeneration
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology
C:Keywords: glycoprotein; hydrolase; kringle; liver; plasma; serine proteinase
F:1-34/Domain: signal sequence #status predicted <SIG>
F:108-148/Domain: fibronectin type II repeat homology <IF2>
F:164-197/Domain: EGF homology <EG1>
F:202-237/Domain: fibronectin type I repeat homology <IF1>
F:245-278/Domain: EGF homology <EG2>
F:286-367/Domain: kringle homology <KRG>

N:Alternate names: plasminogen
C:Species: Capra aegagrus hircus (domestic goat)
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 12-May-1995
C:Accession: C61545
R:Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A:Title: Structural aspects of the plasminogen of various species.
A:Reference number: A61545; MUID:89005015; PMID:3168975
A:Accession: C61545
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-123 <SCH>
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: hydrolase; serine proteinase
F:41-118/Domain: kringle homology <KR4>

Query Match 29.4%; Score 150; DB 2; Length 123;
Best Local Similarity 37.5%; Pred. No. 3e-09;
Matches 33; Conservative 12; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFRGKASTDTMGPCLPWNSATV--LQOTYHAHRSDALQGLGKHNCRNP 60
DB 41 CYHNGQSYRGTSSTVTGRKQCSWSSMIPHRHQKTPESYPNAGLTM-----NYCRNPDA 95
QY 61 RRRPWCYVQVGLKPLV--QECMVHDCAD 86
DB 96 DKSPWCYT---TDPVRWEFCNLKCKSE 120

RESULT 28
A40522
plasmin (EC 3.4.21.7) precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 16-Jul-1999
C:Accession: A40522
R:Kanalas, J.J.; Makker, S.P.
J. Biol. Chem. 266, 10825-10829, 1991
A:Title: Identification of the rat Heymann nephritis autoantigen (GP330) as a receptor
A:Reference number: A40522; MUID:91250378; PMID:1645711
A:Accession: A40522
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-169 <KAN>
A:Cross-references: GB:M62832; NID:G206215; PIDN:AAA41884.1; PID:G554488
A:Note: The authors translated the codon TCT for residue 76 as Ala
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; serine proteinase
F:34-112/Domain: kringle homology <KR4>
F:34-112/55-95,83-107/Disulfide bonds: #status predicted

Query Match 29.3%; Score 149.5; DB 2; Length 169;
Best Local Similarity 37.8%; Pred. No. 4.6e-09;
Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6;

QY 3 CYEGNGHFRGKASTDTMGPCLPWNSATVLTQTYHAHRSDALQ-----GLGKHNCRNP 59
DB 34 CYQNGKSYRGTSSTNTGKQCSW-----VSMTPHSHSKTPANFPDSGL-EMNYCRNP 87
QY 60 N-RRPWCYVQVGLKPLV--QECMVHDCAD 86
DB 88 NDORGWCFT---TDFSRWEYCNLKKCKSE 114

RESULT 29
PLBO
plasmin (EC 3.4.21.7) precursor - bovine
N:Alternate names: plasminogen
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1987 #sequence_revision 28-Apr-1995 #text_change 18-Jun-1999
C:Accession: S45046; A25835; I45961; S03736
R:Berglund, L.; Andersen, M.D.; Petersen, T.E.
submitted to the EMBL Data Library, May 1994
A:Description: Cloning and characterization of the bovine plasminogen cDNA.

A:Reference number: S45046
A:Accession: S45046
A:Molecule type: mRNA
A:Residues: 1-812 <BER>
A:Cross-references: EMBL:X79402; NID:G494962; PIDN:CAAS5939.1; PID:G9494963
A:Experimental source: liver
A:Note: It is uncertain whether Met-1 or Met-8 is the initiator
R:Schaller, J.; Moser, P.W.; Danneberger-Muller, G.A.K.; Rosselet, S.J.; Kampfer, U.; Rickli, J. Biochem. 149, 267-278, 1985
A:Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasminogen
A:Reference number: A25835; MUID:85203906; PMID:3846532
A:Accession: A25835
A:Molecule type: protein
A:Residues: 27-334, 'D', '336-515', 'H', '517-554', 'L', '556-812' <SCH>
R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.
Biochemistry 23, 4243-4250, 1984
A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and bovine plasminogen
A:Reference number: I45961; MUID:85023311; PMID:6148961
A:Accession: I45961
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 706-743, 'R', '745-812' <MAL>
A:Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552
R:Brunscholz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.
Eur. J. Biochem. 114, 465-470, 1981
A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human, bovine and porcine plasminogen
A:Reference number: S03736; MUID:81212097; PMID:7238497
A:Accession: S03736
A:Molecule type: protein
A:Residues: 27-83 <BRU>
C:Function:

A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a variety of tissues; also activates the urokinase-type plasminogen activator
A:Pathway: fibrinolysis
C:Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringle; plasma
F:8-103/Domain: signal sequence #status predicted <SIG>
F:1-26/Domain: plasminogen-related protein precursor homology <PRO>
F:27-812/Product: plasminogen #status experimental <PLPH>
F:27-103/Domain: activation peptide #status experimental <APT>
F:104-583,584-812/Product: plasmin #status experimental <MAT>
F:104-583/Domain: plasmin chain A #status experimental <ACH>
F:110-188/Domain: kringle homology <KR1>
F:192-269/Domain: kringle homology <KR2>
F:282-359/Domain: kringle homology <KR3>
F:384-461/Domain: kringle homology <KR4>
F:485-564/Domain: kringle homology <KR5>
F:584-812/Domain: plasmin chain B #status experimental <BCH>
F:584-805/Domain: trypsin homology <TRY>
F:56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,343-344,345-346,347-348,349-350,351-352,353-354,355-356,357-358,359-360,361-362,363-364,365-366,367-368,369-370,371-372,373-374,375-376,377-378,379-380,381-382,383-384,385-386,387-388,389-390,391-392,393-394,395-396,397-398,399-400,401-402,403-404,405-406,407-408,409-410,411-412,413-414,415-416,417-418,419-420,421-422,423-424,425-426,427-428,429-430,431-432,433-434,435-436,437-438,439-440,441-442,443-444,445-446,447-448,449-450,451-452,453-454,455-456,457-458,459-460,461-462,463-464,465-466,467-468,469-470,471-472,473-474,475-476,477-478,479-480,481-482,483-484,485-486,487-488,489-490,491-492,493-494,495-496,497-498,499-500,501-502,503-504,505-506,507-508,509-510,511-512,513-514,515-516,517-518,519-520,521-522,523-524,525-526,527-528,529-530,531-532,533-534,535-536,537-538,539-540,541-542,543-544,545-546,547-548,549-550,551-552,553-554,555-556,557-558,559-560,561-562,563-564,565-566,567-568,569-570,571-572,573-574,575-576,577-578,579-580,581-582,583-584,585-586,587-588,589-590,591-592,593-594,595-596,597-598,599-600,601-602,603-604,605-606,607-608,609-610,611-612,613-614,615-616,617-618,619-620,621-622,623-624,625-626,627-628,629-630,631-632,633-634,635-636,637-638,639-640,641-642,643-644,645-646,647-648,649-650,651-652,653-654,655-656,657-658,659-660,661-662,663-664,665-666,667-668,669-670,671-672,673-674,675-676,677-678,679-680,681-682,683-684,685-686,687-688,689-690,691-692,693-694,695-696,697-698,699-700,701-702,703-704,705-706,707-708,709-710,711-712,713-714,715-716,717-718,719-720,721-722,723-724,725-726,727-728,729-730,731-732,733-734,735-736,737-738,739-740,741-742,743-744,745-746,747-748,749-750,751-752,753-754,755-756,757-758,759-760,761-762,763-764,765-766,767-768,769-770,771-772,773-774,775-776,777-778,779-780,781-782,783-784,785-786,787-788,789-790,791-792,793-794,795-796,797-798,799-800,801-802,803-804,805-806,807-808,809-810,811-812,813-814,815-816,817-818,819-820,821-822,823-824,825-826,827-828,829-830,831-832,833-834,835-836,837-838,839-840,841-842,843-844,845-846,847-848,849-850,851-852,853-854,855-856,857-858,859-860,861-862,863-864,865-866,867-868,869-870,871-872,873-874,875-876,877-878,879-880,881-882,883-884,885-886,887-888,889-890,891-892,893-894,895-896,897-898,899-900,901-902,903-904,905-906,907-908,909-910,911-912,913-914,915-916,917-918,919-920,921-922,923-924,925-926,927-928,929-930,931-932,933-934,935-936,937-938,939-940,941-942,943-944,945-946,947-948,949-950,951-952,953-954,955-956,957-958,959-960,961-962,963-964,965-966,967-968,969-970,971-972,973-974,975-976,977-978,979-980,981-982,983-984,985-986,987-988,989-990,991-992,993-994,995-996,997-998,999-1000,1001-1002,1003-1004,1005-1006,1007-1008,1009-1010,1011-1012,1013-1014,1015-1016,1017-1018,1019-1020,1021-1022,1023-1024,1025-1026,1027-1028,1029-1030,1031-1032,1033-1034,1035-1036,1037-1038,1039-1040,1041-1042,1043-1044,1045-1046,1047-1048,1049-1050,1051-1052,1053-1054,1055-1056,1057-1058,1059-1060,1061-1062,1063-1064,1065-1066,1067-1068,1069-1070,1071-1072,1073-1074,1075-1076,1077-1078,1079-1080,1081-1082,1083-1084,1085-1086,1087-1088,1089-1090,1091-1092,1093-1094,1095-1096,1097-1098,1099-1100,1101-1102,1103-1104,1105-1106,1107-1108,1109-1110,1111-1112,1113-1114,1115-1116,1117-1118,1119-1120,1121-1122,1123-1124,1125-1126,1127-1128,1129-1130,1131-1132,1133-1134,1135-1136,1137-1138,1139-1140,1141-1142,1143-1144,1145-1146,1147-1148,1149-1150,1151-1152,1153-1154,1155-1156,1157-1158,1159-1160,1161-1162,1163-1164,1165-1166,1167-1168,1169-1170,1171-1172,1173-1174,1175-1176,1177-1178,1179-1180,1181-1182,1183-1184,1185-1186,1187-1188,1189-1190,1191-1192,1193-1194,1195-1196,1197-1198,1199-1200,1201-1202,1203-1204,1205-1206,1207-1208,1209-1210,1211-1212,1213-1214,1215-1216,1217-1218,1219-1220,1221-1222,1223-1224,1225-1226,1227-1228,1229-1230,1231-1232,1233-1234,1235-1236,1237-1238,1239-1240,1241-1242,1243-1244,1245-1246,1247-1248,1249-1250,1251-1252,1253-1254,1255-1256,1257-1258,1259-1260,1261-1262,1263-1264,1265-1266,1267-1268,1269-1270,1271-1272,1273-1274,1275-1276,1277-1278,1279-1280,1281-1282,1283-1284,1285-1286,1287-1288,1289-1290,1291-1292,1293-1294,1295-1296,1297-1298,1299-1300,1301-1302,1303-1304,1305-1306,1307-1308,1309-1310,1311-1312,1313-1314,1315-1316,1317-1318,1319-1320,1321-1322,1323-1324,1325-1326,1327-1328,1329-1330,1331-1332,1333-1334,1335-1336,1337-1338,1339-1340,1341-1342,1343-1344,1345-1346,1347-1348,1349-1350,1351-1352,1353-1354,1355-1356,1357-1358,1359-1360,1361-1362,1363-1364,1365-1366,1367-1368,1369-1370,1371-1372,1373-1374,1375-1376,1377-1378,1379-1380,1381-1382,1383-1384,1385-1386,1387-1388,1389-1390,1391-1392,1393-1394,1395-1396,1397-1398,1399-1400,1401-1402,1403-1404,1405-1406,1407-1408,1409-1410,1411-1412,1413-1414,1415-1416,1417-1418,1419-1420,1421-1422,1423-1424,1425-1426,1427-1428,1429-1430,1431-1432,1433-1434,1435-1436,1437-1438,1439-1440,1441-1442,1443-1444,1445-1446,1447-1448,1449-1450,1451-1452,1453-1454,1455-1456,1457-1458,1459-1460,1461-1462,1463-1464,1465-1466,1467-1468,1469-1470,1471-1472,1473-1474,1475-1476,1477-1478,1479-1480,1481-1482,1483-1484,1485-1486,1487-1488,1489-1490,1491-1492,1493-1494,1495-1496,1497-1498,1499-1500,1501-1502,1503-1504,1505-1506,1507-1508,1509-1510,1511-1512,1513-1514,1515-1516,1517-1518,1519-1520,1521-1522,1523-1524,1525-1526,1527-1528,1529-1530,1531-1532,1533-1534,1535-1536,1537-1538,1539-1540,1541-1542,1543-1544,1545-1546,1547-1548,1549-1550,1551-1552,1553-1554,1555-1556,1557-1558,1559-1560,1561-1562,1563-1564,1565-1566,1567-1568,1569-1570,1571-1572,1573-1574,1575-1576,1577-1578,1579-1580,1581-1582,1583-1584,1585-1586,1587-1588,1589-1590,1591-1592,1593-1594,1595-1596,1597-1598,1599-1600,1601-1602,1603-1604,1605-1606,1607-1608,1609-1610,1611-1612,1613-1614,1615-1616,1617-1618,1619-1620,1621-1622,1623-1624,1625-1626,1627-1628,1629-1630,1631-1632,1633-1634,1635-1636,1637-1638,1639-1640,1641-1642,1643-1644,1645-1646,1647-1648,1649-1650,1651-1652,1653-1654,1655-1656,1657-1658,1659-1660,1661-1662,1663-1664,1665-1666,1667-1668,1669-1670,1671-1672,1673-1674,1675-1676,1677-1678,1679-1680,1681-1682,1683-1684,1685-1686,1687-1688,1689-1690,1691-1692,1693-1694,1695-1696,1697-1698,1699-1700,1701-1702,1703-1704,1705-1706,1707-1708,1709-1710,1711-1712,1713-1714,1715-1716,1717-1718,1719-1720,1721-1722,1723-1724,1725-1726,1727-1728,1729-1730,1731-1732,1733-1734,1735-1736,1737-1738,1739-1740,1741-1742,1743-1744,1745-1746,1747-1748,1749-1750,1751-1752,1753-1754,1755-1756,1757-1758,1759-1760,1761-1762,1763-1764,1765-1766,1767-1768,1769-1770,1771-1772,1773-1774,1775-1776,1777-1778,1779-1780,1781-1782,1783-1784,1785-1786,1787-1788,1789-1790,1791-1792,1793-1794,1795-1796,1797-1798,1799-1800,1801-1802,1803-1804,1805-1806,1807-1808,1809-1810,1811-1812,1813-1814,1815-1816,1817-1818,1819-1820,1821-1822,1823-1824,1825-1826,1827-1828,1829-1830,1831-1832,1833-1834,1835-1836,1837-1838,1839-1840,1841-1842,1843-1844,1845-1846,1847-1848,1849-1850,1851-1852,1853-1854,1855-1856,1857-1858,1859-1860,1861-1862,1863-1864,1865-1866,1867-1868,1869-1870,1871-1872,1873-1874,1875-1876,1877-1878,1879-1880,1881-1882,1883-1884,1885-1886,1887-1888,1889-1890,1891-1892,1893-1894,1895-1896,1897-1898,1899-1900,1901-1902,1903-1904,1905-1906,1907-1908,1909-1910,1911-1912,1913-1914,1915-1916,1917-1918,1919-1920,1921-1922,1923-1924,1925-1926,1927-1928,1929-1930,1931-1932,1933-1934,1935-1936,1937-1938,1939-1940,1941-1942,1943-1944,1945-1946,1947-1948,1949-1950,1951-1952,1953-1954,1955-1956,1957-1958,1959-1960,1961-1962,1963-1964,1965-1966,1967-1968,1969-1970,1971-1972,1973-1974,1975-1976,1977-1978,1979-1980,1981-1982,1983-1984,1985-1986,1987-1988,1989-1990,1991-1992,1993-1994,1995-1996,1997-1998,1999-2000,2001-2002,2003-2004,2005-2006,2007-2008,2009-2010,2011-2012,2013-2014,2015-2016,2017-2018,2019-2020,2021-2022,2023-2024,2025-2026,2027-2028,2029-2030,2031-2032,2033-2034,2035-2036,2037-2038,2039-2040,2041-2042,2043-2044,2045-2046,2047-2048,2049-2050,2051-2052,2053-2054,2055-2056,2057-2058,2059-2060,2061-2062,2063-2064,2065-2066,2067-2068,2069-2070,2071-2072,2073-2074,2075-2076,2077-2078,2079-2080,2081-2082,2083-2084,2085-2086,2087-2088,2089-2090,2091-2092,2093-2094,2095-2096,2097-2098,2099-2100,2101-2102,2103-2104,2105-2106,2107-2108,2109-2110,2111-2112,2113-2114,2115-2116,2117-2118,2119-2120,2121-2122,2123-2124,2125-2126,2127-2128,2129-2130,2131-2132,2133-2134,2135-2136,2137-2138,2139-2140,2141-2142,2143-2144,2145-2146,2147-2148,2149-2150,2151-2152,2153-2154,2155-2156,2157-2158,2159-2160,2161-2162,2163-2164,2165-2166,2167-2168,2169-2170,2171-2172,2173-2174,2175-2176,2177-2178,2179-2180,2181-2182,2183-2184,2185-2186,2187-2188,2189-2190,2191-2192,2193-2194,2195-2196,2197-2198,2199-2200,2201-2202,2203-2204,2205-2206,2207-2208,2209-2210,2211-2212,2213-2214,2215-2216,2217-2218,2219-2220,2221-2222,2223-2224,2225-2226,2227-2228,2229-2230,2231-2232,2233-2234,2235-2236,2237-2238,2239-2240,2241-2242,2243-2244,2245-2246,2247-2248,2249-2250,2251-2252,2253-2254,2255-2256,2257-2258,2259-2260,2261-2262,2263-2264,2265-2266,2267-2268,2269-2270,2271-2272,2273-2274,2275-2276,2277-2278,2279-2280,2281-2282,2283-2284,2285-2286,2287-2288,2289-2290,2291-2292,2293-2294,2295-2296,2297-2298,2299-2300,2301-2302,2303-2304,2305-2306,2307-2308,2309-2310,2311-2312,2313-2314,2315-2316,2317-2318,2319-2320,2321-2322,2323-2324,2325-2326,2327-2328,2329-2330,2331-2332,2333-2334,2335-2336,2337-2338,2339-2340,2341-2342,2343-2344,2345-2346,2347-2348,2349-2350,2351-2352,2353-2354,2355-2356,2357-2358,2359-2360,2361-2362,2363-2364,2365-2366,2367-2368,2369-237

Search completed: December 3, 2003, 14:44:12
Job time : 4.55172 secs

Result No.	Score	Match	Query %	DB	ID	Description
1	51.0	100.0	431	1	UROK_HUMAN	P00749 homo sapien
2	47.5	93.1	433	1	UROK_PAPCY	P16227 papio cynoc
3	42.0	82.4	432	1	UROK_PIG	P04185 sus scrofa
4	40.6	79.6	442	1	UROK_RAT	P29598 rattus norv
5	39.6	77.6	433	1	UROK_BOVIN	P05589 bos taurus
6	38.7	75.9	433	1	UROK_MOUSE	P06869 mus musculu
7	24.1	47.3	431	1	URTB_DSRO	P98121 desmodus ro
8	24.1	47.3	477	1	URTB_DSRO	P15638 desmodus ro
9	22.6	44.3	562	1	TPA_HUMAN	P00750 homo sapien
10	22.5	44.1	559	1	TPA_RAT	P19637 rattus norv
11	22.0	43.1	394	1	URTG_DSRO	P49150 desmodus ro
12	21.9	42.9	559	1	TPA_MOUSE	P11214 mus musculu
13	21.3	41.8	477	1	URTI_DSRO	P98119 desmodus ro
14	20.9	41.0	566	1	UROK_BOVIN	Q28198 bos taurus
15	198.5	38.9	434	1	UROK_CHICK	P15120 gallus gail
16	193	37.8	655	1	HGFA_HUMAN	Q04756 homo sapien
17	188	36.9	653	1	HGFA_MOUSE	Q09098 mus musculu
18	180.5	35.4	603	1	FA12_CAVPO	Q04962 cavia porce
19	170.5	33.4	615	1	FA12_HUMAN	P00748 homo sapien
20	161	31.6	1420	1	APOA_MACMU	P14417 macaca mula
21	159	31.2	473	1	KRMI_MOUSE	Q99n43 mus musculu
22	159	31.2	473	1	KRMI_RAT	Q924s4 rattus norv
23	158	31.0	452	1	KRMI_XENLA	Q90y90 xenopus lae
24	156	30.6	475	1	KRMI_HUMAN	Q96mu8 homo sapien
25	156	30.6	458	1	APOA_HUMAN	P08519 homo sapien
26	155.5	30.5	593	1	FA12_BOVIN	P98140 bos taurus
27	149.5	29.3	169	1	PLMN_RAT	Q01177 rattus norv
28	149	29.2	812	1	PLMN_BOVIN	P06868 bos taurus
29	148.5	29.1	462	1	KRM2_HUMAN	Q8ncw0 homo sapien
30	148.5	29.1	810	1	PLMN_ERIEU	Q29485 erinaceus e
31	148	29.0	790	1	PLMN_PIG	P06867 sus scrofa
32	147.5	28.9	943	1	ROR2_HUMAN	Q01974 homo sapien
33	147.5	28.9	944	1	ROR2_MOUSE	Q9a138 mus musculu

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carncini P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Touchman J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kraywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[7]
PP SEQUENCE OF 66-431 FROM N.A.
RX MEDLINE=84272706; PubMed=6589620;
RA Verde P., Stoppelli M.P., Galeffi P., di Nocera P., Blasi F.;
RT "Identification and primary sequence of an unspliced human urokinase
RT poly(A)+ RNA.";
Proc. Natl. Acad. Sci. U.S.A. 81:4727-4731(1984).
[8]
PP SEQUENCE OF 21-177.
RX MEDLINE=83055084; PubMed=6754569;
RA Gunzler W.A., Steffens G.J., Otting F., Kim S.-M.A., Frankus E.,
RA Flohe L.;
RT "The primary structure of high molecular mass urokinase from human
RT urine. The complete amino acid sequence of the A chain.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1155-1165(1982).
[9]
PP SEQUENCE OF 156-176 AND 179-224.
RX MEDLINE=83003608; PubMed=6749491;
RA Schaller J., Nick H., Rickli E.E., Gillesen D., Iergier W.,
RA Studer R.O.;
RT "Human low-molecular-weight urinary urokinase. Partial
RT characterization and preliminary sequence data of the two polypeptide
RT chains.";
RL Eur. J. Biochem. 125:251-257(1982).
[10]
PP SEQUENCE OF 158-410.
RX MEDLINE=83055099; PubMed=6754572;
RA Steffens G.J., Gunzler W.A., Otting F., Frankus E., Flohe L.;
RT "The complete amino acid sequence of low molecular mass urokinase
RT from human urine.";
RL Hoppe-Seyler's Z. Physiol. Chem. 363:1043-1058(1982).
[11]
PP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=96000858; PubMed=8591045;
RA Spraggon G., Phillips C., Nowak U.K., Ponting C.P., Saunders D.,
RA Dobson C.M., Stuart D.I., Jones E.Y.;
RT "The crystal structure of the catalytic domain of human
RL urokinase-type plasminogen activator.";
Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
[12]
PP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 159-411.
RX MEDLINE=20266327; PubMed=10805774;
RA Sperl S., Jacob U., Arroyo de Prada N., Sturzebecher J., Wilhelm O.G.,
RA Bode W., Magdolen V., Huber R., Moroder L.;
RT "(4-aminomethyl)phenylguanidine derivatives as nonpeptidic highly
RT selective inhibitors of human urokinase.";
Proc. Natl. Acad. Sci. U.S.A. 97:5113-5118(2000).
[13]
PP STRUCTURE BY NMR.
RX MEDLINE=89127526; PubMed=2536903;
RA Oswald R.E., Bogusky M.J., Bamberger M., Smith R.A.G., Dobson C.M.;
RT "Dynamics of the multidomain fibrinolytic protein urokinase from two-
RT dimensional NMR.";
RL Nature 337:579-582(1989).
[14]
PP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=93003110; PubMed=1327118;
RA Li X., Smith R.A.G., Dobson C.M.;
RT "Sequential 1H NMR assignments and secondary structure of the kringle
RT domain from urokinase.";
Biochemistry 31:9562-9571(1992).
[15]
PP STRUCTURE BY NMR OF 67-155.
RX MEDLINE=94149701; PubMed=8107091;
RA Li X., Bokman A.M., Llinas M., Smith R.A.G., Dobson C.M.;
RT "Solution structure of the kringle domain from urokinase-type
RT plasminogen activator.";
J. Mol. Biol. 235:1548-1559(1994).
[16]
PP VARIANT LEU-141.
RX MEDLINE=96186279; PubMed=8652631;
RA Yoshimoto M., Ushiyama Y., Sakai M., Tanaki S., Hara H., Takahashi K.,
RA Sawasaki Y., Hanada K.;
RT "Characterization of single chain urokinase-type plasminogen
RT activator with a novel amino-acid substitution in the kringle
RT structure.";
Biochim. Biophys. Acta 1293:83-89(1996).
[17]
PP VARIANT LEU-141.
RX MEDLINE=97218551; PubMed=9065988;
RA Conne B., Berczy M., Belin D.;
RT "Detection of polymorphisms in the human urokinase-type plasminogen
RT activator gene.";
Thromb. Haemost. 77:434-435(1997).
[18]
PP ERRATUM.
RA Conne B., Berczy M., Belin D.;
RL Thromb. Haemost. 78:973-973(1997).
[19]
PP VARIANT LEU-141.
RX MEDLINE=973737920; PubMed=9194591;
RA Turkmen B., Schmitt M., Schmalfeidt B., Trommler P., Hell W.,
RA Creutzburg S., Graeff H., Magdolen V.;
RT "Mutational analysis of the genes encoding urokinase-type plasminogen
RT activator (uPA) and its inhibitor PAI-1 in advanced ovarian cancer.";
Electrophoresis 18:686-689(1997).
CC -!- FUNCTION: POTENT PLASMINOGEN ACTIVATOR AND IS CLINICALLY USED FOR
CC THERAPY OF THROMBOLYTIC DISORDERS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN.
CC -!- PHARMACEUTICAL: Available under the name Abbokinase (Abbott). Used
CC in Pulmonary Embolism (PE) to initiate fibrinolysis.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X02419; CAA26268.1; -
DR EMBL; M15476; AAA61253.1; -
DR EMBL; D00244; BAA00175.1; -
DR EMBL; D11143; BAA01919.1; -
DR EMBL; X02760; CAA26535.1; -
DR EMBL; AF377330; AAK53822.1; -
DR EMBL; BC013575; AAH13575.1; -
DR EMBL; K03226; AAC97138.1; -
DR EMBL; K02286; AAA61252.1; -
DR EMBL; A21571; CAA01559.1; -
DR EMBL; A18397; CAA01390.1; -
DR PIR; A00931; UKHU.
PDB; 1XDU; 31-OCT-93.

Query Match 100.0%; Score 510; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.6e-51;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 68 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 128 RRRPWCYVQVGLKPLVQECMVHDCADGK 155

RESULT 2
ID UROK PAPCY STANDARD; PRT; 433 AA.
AC P16227;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAÜ.
OS Papio cynocephalus (Yellow baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9556;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thoracic aorta;
RX MEDLINE=90287734; PubMed=2113276;
RA Au Y.P.T., Wang T.W., Clowes A.W.;
RT "Nucleotide and deduced amino acid sequences of baboon urokinase-type plasminogen activator."
RL Nucleic Acids Res. 18:3411-3411(1990).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 155 IN THE LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51935; CAA36200.1; -;
DR PUR; S14587; UKBAY.
DR HSP; P00749; ILMW.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.

DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 176 CHAIN A (BY SIMILARITY).
FT CHAIN 155 176 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 178 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 26 62 EGF-LIKE.
FT DOMAIN 69 150 KRINGLE.
FT DOMAIN 151 177 CONNECTING PEPTIDE.
FT DOMAIN 178 433 SERINE PROTEASE.
FT DISULFID 30 38 BY SIMILARITY.
FT DISULFID 32 50 BY SIMILARITY.
FT DISULFID 52 61 BY SIMILARITY.
FT DISULFID 167 298 INTERCHAIN (BY SIMILARITY).
FT DISULFID 208 224 BY SIMILARITY.
FT DISULFID 216 287 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 223 223 CHARGE RELAY SYSTEM.
FT ACT_SITE 274 274 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
FT CARBOHYD 324 324 N-LINKED (GLCNAC...) (BY SIMILARITY).
SQ SEQUENCE 433 AA; 48595 MW; 816D22DFEDDC8792 CRC64;

Query Match 93.1%; Score 475; DB 1; Length 433;
Best Local Similarity 94.3%; Pred. No. 1.8e-47;
Matches 83; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 60
DB 67 KTCYEGNGHFRGKASTDTMGRCLPWNATVLOQTYHAHRSALQLGLGKHNYCRNPDN 126

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 127 RRRPWCYVQVGLKPLVQECMVHDCADGK 154

RESULT 3
UROK PIG
ID UROK PIG STANDARD; PRT; 442 AA.
AC P04185;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAÜ.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
EX MEDLINE=85087954; PubMed=6096832;
RA Nagamine Y., Pearson D., Altus M.S., Reich E.;
RT "cDNA and gene nucleotide sequence of porcine plasminogen activator."
RL Nucleic Acids Res. 12:9525-9541(1984).
RN [2]
RP REVISION TO 241.
RA Nagamine Y.;
RL Submitted (DEC-1986) to the PIR data bank.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

```
CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X01648; CAA25806.1; -.
CC DR EMBL; X02724; CAA26511.1; -.
CC DR PIR; A00932; UKPG.
CC DR HSSP; P00749; LKDU.
CC DR MEROPS; S01.231; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR000001; Kringle.
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR Pfam; PF00051; kringle; 1.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR PRINTS; PR00018; KRINGLE.
CC DR ProDom; PD000395; Kringle; 1.
CC DR SMART; SM00130; KR; 1.
CC DR SMART; SM00020; Tryp_Spc; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS01186; EGF_2; FALSE NEG.
CC DR PROSITE; PS00021; KRINGLE_1; 1.
CC DR PROSITE; PS00070; KRINGLE_2; 1.
CC DR PROSITE; PS50240; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20
FT CHAIN 21 442 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 188 CHAIN A (BY SIMILARITY).
FT CHAIN 190 442 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 189 CONNECTING PEPTIDE.
FT DOMAIN 190 442 SERINE PROTEASE.
FT CARBOHYD 152 152 N-LINKED (GLCNAc. . .).
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 INTERCHAIN (BY SIMILARITY).
FT DISULFID 179 310 BY SIMILARITY.
FT DISULFID 220 236 BY SIMILARITY.
FT DISULFID 228 299 BY SIMILARITY.
FT DISULFID 324 393 BY SIMILARITY.
FT DISULFID 356 372 BY SIMILARITY.
FT DISULFID 383 411 BY SIMILARITY.
FT ACT_SITE 235 235 CHARGE RELAY SYSTEM.
FT ACT_SITE 286 286 CHARGE RELAY SYSTEM.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM.
FT CONFLICT 241 241 Q -> H (IN REF. 1; CAA25806).
FT CONFLICT 242 242 Q -> H (IN REF. 1; CAA25811).
FT CONFLICT 288 288 A -> GS (IN REF. 1; CAA25806).
SQ SEQUENCE 442 AA; 49116 MW; EE32PCEFS01321EE CRC64;
Query Match 82.4%; Score 420; DB 1; Length 442;
Best Local Similarity 81.8%; Pred. No. 4,1e-41;
Matches 72; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 1 KTCVEGNHGYFGKASTDMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 QTCFEGNHSYRGKANTNGRCPCLPWSATVLLNTYHAHSPDALQLGLGKHNYCRNPDN 129
QY 61 RRRPKCYVQGLKPLVQECMWHDCADGK 88
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
130 QRRPKCYVQVGLKQLVQECMVPNCSGGE 157
```

```
RESULT 4
UROK_RAT
ID UROK_RAT STANDARD; PRT; 432 AA.
AC P29598;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAUF.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer 344;
RX MEDLINE=92233409; PubMed=1568219;
RA Henderson B.R.; Tansey W.P.; Phillips S.M.; Ramshaw I.A.;
RA Keiford R.F.;
RT "Transcriptional and posttranscriptional activation of urokinase
RT plasminogen activator gene expression in metastatic tumor cells.";
RL Cancer Res. 52:2489-2496(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Rabbani S.A.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH
CC CONSISTS OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM
CC CONTAINS A LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE
CC LOW MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC DR EMBL; X63434; CAA45028.1; -.
CC DR EMBL; X65651; CAA46601.1; -.
CC DR PIR; S24604; S18932.
CC DR HSSP; P00749; LKDU.
CC DR MEROPS; S01.231; -.
CC DR InterPro; IPR001314; Chymotrypsin.
CC DR InterPro; IPR006209; EGF like.
CC DR InterPro; IPR006210; IEGF.
CC DR InterPro; IPR000001; Kringle.
CC DR InterPro; IPR001254; Ser_protease_Try.
CC DR Pfam; PF00051; kringle; 1.
CC DR Pfam; PF00089; trypsin; 1.
CC DR PRINTS; PR00722; CHYMOTRYPSIN.
CC DR PRINTS; PR00018; KRINGLE.
CC DR ProDom; PD000395; Kringle; 1.
CC DR SMART; SM00130; KR; 1.
CC DR SMART; SM00020; Tryp_Spc; 1.
CC DR PROSITE; PS00022; EGF_1; 1.
CC DR PROSITE; PS01186; EGF_2; FALSE NEG.
CC DR PROSITE; PS00021; KRINGLE_1; 1.
CC DR PROSITE; PS00021; KRINGLE_2; 1.
CC DR PROSITE; PS50070; TRYPSIN_DOM; 1.
CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
```

```

KW Plasminogen activation: Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 432 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 20 432 CHAIN A (BY SIMILARITY).
FT CHAIN 156 177 SHORT A CHAIN (A1) (BY SIMILARITY).
FT CHAIN 179 432 CHAIN B (BY SIMILARITY).
FT DOMAIN 27 63 EGF-LIKE.
FT DOMAIN 70 151 KRINGLE.
FT DOMAIN 152 178 CONNECTING PEPTIDE.
FT DOMAIN 179 432 SERINE PROTEASE.
FT DISULFID 31 39 BY SIMILARITY.
FT DISULFID 33 51 BY SIMILARITY.
FT DISULFID 53 62 BY SIMILARITY.
FT DISULFID 168 300 INTERCHAIN (BY SIMILARITY).
FT DISULFID 210 226 BY SIMILARITY.
FT DISULFID 218 289 BY SIMILARITY.
FT DISULFID 314 383 BY SIMILARITY.
FT DISULFID 346 362 BY SIMILARITY.
FT DISULFID 373 401 BY SIMILARITY.
FT ACT_SITE 225 225 CHARGE RELAY SYSTEM.
FT ACT_SITE 276 276 CHARGE RELAY SYSTEM.
FT ACT_SITE 377 377 CHARGE RELAY SYSTEM.
FT CONFLICT 16 16 N -> H (IN REF. 2).
FT CONFLICT 24 24 E -> G (IN REF. 2).
FT CONFLICT 332 332 D -> N (IN REF. 2).
SQ SEQUENCE 432 AA; 47957 MW; 4EB1B96C716244C8 CRC64;

Query Match 79.6%; Score 406; DB 1; Length 432;
Best Local Similarity 79.5%; Pred. No. 1.7e-39;
Matches 70; Conservative 5; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDN 60
DB 68 KTCYQGNHGYRGKANDTDTGRPCLPWNSPAVLQOTYNAHRSDALSLGLGKHNYCRNPDN 127

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 128 QRRPWCYVQIGLKQFVQECMVQDCSLK 155

RESULT 5
UROK_BOVIN STANDARD; PRT; 433 AA.
AC Q05589; Q28209;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLA0.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP Schiencing W.-D.;
RC TISSUE=Aortic endothelium;
RX MEDLINE=93216119; PubMed=8385052;
RA Kraetzschmar J., Haendler B., Kojima S., Rifkin D.B.,
RA Schiencing W.-D.;
RT "Bovine urokinase-type plasminogen activator and its receptor:
RT cloning and induction by retinoic acid.";
PL Gene 125:177-183(1993).
RP [2]
RW SEQUENCE OF 12-433 FROM N.A.
RC TISSUE=Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RT and tPA.";
RL Int. Dairy J. 5:605-617(1995).
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in

```

```

CC plasminogen to form plasmin.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L03546; AAA51419.1; -.
CC EMBL; X85801; CAA59796.1; -.
CC PIR; JN0560; JN0560.
CC HSP; P00749; 1LMW.
CC MEROPS; S01.231; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR001254; Ser_Protease_Try.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00020; Tryp_Spc; 1.
CC PROSITE; PS00022; EGF_1; 1.
CC PROSITE; PS01186; EGF_2; FALSE NEG.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00070; KRINGLE_2; 1.
CC PROSITE; PS02040; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 179 CHAIN A (BY SIMILARITY).
FT CHAIN 181 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 29 65 EGF-LIKE.
FT DOMAIN 72 153 KRINGLE.
FT DOMAIN 154 180 CONNECTING PEPTIDE.
FT DOMAIN 181 433 SERINE PROTEASE.
FT DISULFID 33 41 BY SIMILARITY.
FT DISULFID 35 53 BY SIMILARITY.
FT DISULFID 55 64 BY SIMILARITY.
FT DISULFID 170 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 189 189 A -> T (IN REF. 2).
SQ SEQUENCE 433 AA; 48730 MW; 4DE1B8D4DA47027A CRC64;

Query Match 77.6%; Score 396; DB 1; Length 433;
Best Local Similarity 76.1%; Pred. No. 2.4e-36;
Matches 67; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOOTYHAHRSDALQGLGKHNYCRNPDN 60
DB 70 KTCYQGNHGYRGKANDTDTGRPCLPWNSPAVLQOTYNAHRSDALSLGLGKHNYCRNPDN 129

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 130 QRRPWCYVQIGLKQFVQECMVQDCSVK 157

```

```
RESULT 6
ID UROK MOUSE STANDARD; PRT; 433 AA.
AC P06869;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator).
GN PLAU
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP MEDLINE=85179474; PubMed=2985383;
RA Belin D., Vassalli J.-D., Conbepine C., Godeau F., Nagamine Y.,
RA Reich E., Kocher H.P., Duvoisin R.M.;
RT "Cloning, nucleotide sequencing and expression of cDNAs encoding
RT mouse urokinase-type plasminogen activator.";
RL Eur. J. Biochem. 148:225-232 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RP MEDLINE=88163489; PubMed=2831940;
RA Degen S.J.F., Heckel J.L., Reich E., Degen J.L.;
RT "The murine urokinase-type plasminogen activator gene.";
RL Biochemistry 26:8270-8279 (1987).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SUBUNIT: FOUND IN HIGH AND LOW MOLECULAR MASS FORMS. EACH CONSISTS
CC OF TWO CHAINS, A AND B. THE HIGH MOLECULAR MASS FORM CONTAINS A
CC LONG CHAIN A. CLEAVAGE OCCURS AFTER RESIDUE 156 IN THE LOW
CC MOLECULAR MASS FORM TO YIELD A SHORT A1 CHAIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR ENBL; X02389; CAA26231.1; -.
DR EMBL; M17922; AAA40539.1; -.
DR PIR; A29420; UKMS.
DR HSP; P00749; IKDU.
DR MEROPS; S01.231; -.
DR MGD; MGI:97611; Plau.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; FALSE NEG.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
```

```
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Zymogen; Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 433 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 178 CHAIN A (BY SIMILARITY).
FT CHAIN 157 178 SHORT A CHAIN (A1)
FT CHAIN 180 433 CHAIN B (BY SIMILARITY).
FT DOMAIN 28 64 EGF-LIKE.
FT DOMAIN 71 152 KRINGLE.
FT DOMAIN 153 179 CONNECTING PEPTIDE.
FT DOMAIN 180 433 SERINE PROTEASE.
FT DISULFID 32 40 BY SIMILARITY.
FT DISULFID 34 52 BY SIMILARITY.
FT DISULFID 54 63 BY SIMILARITY.
FT DISULFID 169 301 INTERCHAIN (BY SIMILARITY).
FT DISULFID 211 227 BY SIMILARITY.
FT DISULFID 219 290 BY SIMILARITY.
FT DISULFID 315 384 BY SIMILARITY.
FT DISULFID 347 363 BY SIMILARITY.
FT DISULFID 374 402 BY SIMILARITY.
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM.
FT ACT_SITE 277 277 CHARGE RELAY SYSTEM.
FT ACT_SITE 378 378 CHARGE RELAY SYSTEM.
SQ SEQUENCE 433 AA; 48268 MW; A99C35F6250443F9 CRC64;

Query Match 75.9%; Score 387; DB 1; Length 433;
Best Local Similarity 72.7%; Pred. No. 2.6e-37;
Matches 64; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

Qy 1 KTCVGNHGYRCKASTDTMGRCPLWNSTVLQOTVHAHRSALQGLGKHNYCRNPDN 60
Db 69 KTCYHGNGSYRGKANTDTKGRCLAWNAPVLQKPYNAHRPDAISLGKHKNYCRNPDN 128
Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 129 QKRPWCYVQIGLRQFVQECMVHDCSLK 156

RESULT 7
URTB DESRO STANDARD; PRT; 431 AA.
AC P98121;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator beta precursor (EC 3.4.21.68) (DSPA
DE beta).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Salivary gland;
RP MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237 (1991).
RN [2]
RP CHARACTERIZATION.
RP MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
RA Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
RT vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403 (1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
CC EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
```

AGENT.

-1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in plasminogen to form plasmin.

-1- SUBUNIT: Monomer.

-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

-1- SIMILARITY: Contains 1 kringle domain.

-1- SIMILARITY: Contains 1 EGF-like domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M63989; AAA31594.1; --

PIR; JS0599; JS0599.

HSP; P98119; IAS1.

MEROPS; S01.239; --

InterPro; IPR001314; Chymotrypsin.

InterPro; IPR006209; EGF like.

InterPro; IPR006210; IEFG.

InterPro; IPR000001; Kringle.

InterPro; IPR001254; Ser_Protease_Try.

Pfam; PF00008; EGF; 1.

Pfam; PF00051; kringle; 1.

Pfam; PF00089; trypsin; 1.

PRINTS; PR00722; CHYMOTRYPSIN.

PRINTS; PR00018; KRINGLE.

ProDom; PD000395; Kringle; 1.

SMART; SM00181; EGF; 1.

SMART; SM00130; KG; 1.

SMART; SM00020; Tryp_SPC; 1.

PROSITE; PS00022; EGF_1; 1.

PROSITE; PS01186; EGF_2; 1.

PROSITE; PS00021; KRINGLE_1; 1.

PROSITE; PS00070; KRINGLE_2; 1.

PROSITE; PS00240; TRYPSIN_DOM; 1.

PROSITE; PS00134; TRYPSIN_HIS; 1.

PROSITE; PS00135; TRYPSIN_SER; 1.

Plasminogen activation; Hydrolase; Serine protease; Glycoprotein; Kringle; EGF-like domain; Signal; Multigene family.

SIGNAL 1 36 POTENTIAL.

CHAIN 37 431 SALIVARY PLASMINOGEN ACTIVATOR BETA.

DOMAIN 37 75 EGF-LIKE.

DOMAIN 82 163 KRINGLE.

DOMAIN 179 431 SERINE PROTEASE.

ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT_SITE 275 275 CHARGE RELAY SYSTEM (BY SIMILARITY).

ACT_SITE 382 382 CHARGE RELAY SYSTEM (BY SIMILARITY).

DISELFID 41 52 BY SIMILARITY.

DISELFID 46 63 BY SIMILARITY.

DISELFID 65 74 BY SIMILARITY.

DISELFID 82 163 BY SIMILARITY.

DISELFID 103 145 BY SIMILARITY.

DISELFID 134 158 BY SIMILARITY.

DISELFID 168 299 BY SIMILARITY.

DISELFID 211 227 BY SIMILARITY.

DISELFID 219 288 BY SIMILARITY.

DISELFID 313 388 BY SIMILARITY.

DISELFID 345 361 BY SIMILARITY.

DISELFID 378 406 BY SIMILARITY.

CARBOHYD 139 139 N-LINKED (GLCNAC...) (POTENTIAL).

CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).

SEQUENCE 431 AA; 48221 MW; 699B5E75B162CBF CRC64;

Query Match 47.3%; Score 241; DB 1; Length 431;
Best Local Similarity 50.0%; Pred.No.2e-20;
Matches 42; Conservative 12; Mismatches 30; Indels 0; Gaps 0;

2 TCYKGGHVFYRGKASDTWGRCLPWNSATLVQTYHAHRSDAITLGHNHCNPDNR 61
||||| |||| | : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 81 TCYKGGVTVRGTWTSSEGAQCIWNNSMLLTRTYNGRRSDAITLGHNHCNPDNR 144
QY 62 RRPWCYVQVLKPLVOECMVHDCA 85
DB 141 SKWCVCVIKASKFILEFCSPVCS 164

RESULT 8
URT2 DESRO
ID URT2_DESGRO STANDARD; PRT; 477 AA.
AC P15638;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 2 precursor (EC 3.4.21.68) (DSFA
alpha-2) (BAT-PA) (T-plasminogen activator).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
Alagon A., Donner P., Schleuning W.D.;
RT "The plasminogen activator family from the salivary gland of the
vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Salivary gland;
RX MEDLINE=90036867; PubMed=2509450;
RA Gardell S.J., Duong L.T., Diehl R.E., York J.D., Hare T.R.,
Register R.B., Jacobs J.W., Dixon K.A.F., Friedman P.A.;
RT "Isolation, characterization, and cDNA cloning of a vampire bat
salivary plasminogen activator.";
RL J. Biol. Chem. 264:17947-17952(1989).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,
Donner P.;
RT "Plasminogen activators from the saliva of Desmodus rotundus (common
vampire bat): unique fibrin specificity.";
RL Ann. N.Y. Acad. Sci. 667:395-403(1992).
CC -1- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. PROBABLE POTENT THROMBOLYTIC
AGENT.
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
plasminogen to form plasmin.
CC -1- ENZYME REGULATION: ACTIVITY TOWARD PLASMINOGEN IS STIMULATED IN
THE PRESENCE OF FIBRIN I.
CC -1- SUBUNIT: Monomer.
CC -1- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
STIMULATION OF ACTIVITY.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; M63988; AAA31593.1; --

ID TPA RAT STANDARD; PRT; 559 AA.
 AC P19637;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
 DE (t-PA) (t-plasminogen activator).
 GN PLAT.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_taxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. PubMed=3148445;
 RX MEDLINE=90130448; PubMed=2105315;
 RA Feng P., Ohlsson M., Ny T.;
 RT "The structure of the TAT-less rat tissue-type plasminogen activator
 gene. Species-specific sequence divergences in the promoter predict
 differences in regulation of gene expression.";
 RL J. Biol. Chem. 265:2022-2027(1990).
 CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
 TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
 CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
 ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
 MANY OTHER PHYSIOLOGICAL EVENTS.
 CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
 plasminogen to form plasmin.
 CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
 BOND.
 CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
 CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
 PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
 ARG-308 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
 CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
 CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 EGF-like domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 2 kringle domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M23697; AAA41812.1; -;
 DR EMBL; M31197; AAA42261.1; -;
 DR EMBL; M31185; AAA42261.1; JOINED.
 DR EMBL; M31186; AAA42261.1; JOINED.
 DR EMBL; M31187; AAA42261.1; JOINED.
 DR EMBL; M31188; AAA42261.1; JOINED.
 DR EMBL; M31189; AAA42261.1; JOINED.
 DR EMBL; M31190; AAA42261.1; JOINED.
 DR EMBL; M31191; AAA42261.1; JOINED.
 DR EMBL; M31192; AAA42261.1; JOINED.
 DR EMBL; M31193; AAA42261.1; JOINED.
 DR EMBL; M31194; AAA42261.1; JOINED.
 DR EMBL; M31195; AAA42261.1; JOINED.
 DR EMBL; M31196; AAA42261.1; JOINED.
 DR EMBL; A19618; AAA01482.1; -;
 DR PIR; A35029; A35029.
 DR HGSP; P00750; IRTF.
 DR MEROPS; S01.232; -.

DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibrinctn1.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00039; fnl; 1.
 DR Pfam; PF00051; kringle; 2.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000395; Kringle; 2.
 DR SMART; SM00181; EGF; 1.
 DR SMART; SM00058; FNI; 1.
 DR SMART; SM00130; KR; 2.
 DR SMART; SM00020; Tryp Spc; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00021; KRINGLE_1; 2.
 DR PROSITE; PS00070; KRINGLE_2; 2.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
 KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
 FT SIGNAL 1 17
 FT PROPEP 18 29
 FT CHAIN 30 559
 FT CHAIN 30 308
 FT CHAIN 309 559
 FT CHAIN 36 78
 FT DOMAIN 79 117
 FT DOMAIN 124 205
 FT DOMAIN 213 294
 FT DOMAIN 309 559
 FT ACT SITE 355 355
 FT ACT SITE 404 404
 FT ACT SITE 510 510
 FT DISULFID 38 68
 FT DISULFID 66 75
 FT DISULFID 83 94
 FT DISULFID 88 105
 FT DISULFID 107 116
 FT DISULFID 124 205
 FT DISULFID 145 187
 FT DISULFID 176 200
 FT DISULFID 213 294
 FT DISULFID 234 276
 FT DISULFID 265 289
 FT DISULFID 297 428
 FT DISULFID 340 356
 FT DISULFID 348 417
 FT DISULFID 442 516
 FT DISULFID 474 490
 FT DISULFID 506 534
 FT CARBOHYD 149 149
 FT CARBOHYD 481 481
 FT CONFLICT 380 380
 SQ SEQUENCE 559 AA; 62903 MW; 7DBD3809C1DC921 CRC64;

Query Match 44.1%; Score 225; DB 1; Length 559;
 Best Local Similarity 47.7%; Pred. No. 1.9e-18;
 Matches 41; Conservative 10; Mismatches 35; Indels 0; Gaps 0;

Qy 2 TCVEGNCHFRGKASTMTGRPCLPWNSATVLQTYHAHRS DALQLGLGKHNYCRNPDNR 61
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 123 TCPEGGQITRYGTWMTAENGAEICINWNSSALSKQPSYARRPNAIKLGLGHNYCRNPDNR 182
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Qy 62 RRPWCYVQVGLKPLVQECVHDCADG 87

KW	Cringle; Signal; Multigene family.
FT	SIGNAL 1 36 POTENTIAL.
FT	CHAIN 37 394 SALIVARY PLASMINOGEN ACTIVATOR GAMMA.
FT	DOMAIN 45 126 KRINGLE.
FT	DOMAIN 142 394 SERINE PROTEASE.
FT	ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 238 238 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE 345 345 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	DISULFID 45 126 BY SIMILARITY.
FT	DISULFID 66 108 BY SIMILARITY.
FT	DISULFID 97 121 BY SIMILARITY.
FT	DISULFID 131 262 BY SIMILARITY.
FT	DISULFID 174 190 BY SIMILARITY.
FT	DISULFID 182 251 BY SIMILARITY.
FT	DISULFID 276 351 BY SIMILARITY.
FT	DISULFID 308 324 BY SIMILARITY.
FT	DISULFID 341 369 BY SIMILARITY.
FT	CARBOHYD 315 315 N-LINKED (GLCNAC...) (POTENTIAL).
SO	SEQUENCE 394 AA; 44105 MW; 9CCD6F52F3D81FCD CRC64;
Query Match	
Best Local Similarity 43.1%; Score 220; DB 1; Length 394;	
Matches 38; Conservative 13; Mismatches 33; Indels 0; Gaps 0;	
Qy	2 TCYEGNGHFYRGKASTDTMGRPCLPNSATVLTQTYHAHRSDALQIGLGKHNVCRNPDNR 61
Db	44 TCYKQDQVTYRGWTSTSESGAQCIINNNSLLIRRTYNGRMPEAVKVLGLGNHNYCRNPDGA 103
Qy	62 RRPWCYVOVGLKPLVQECMYHDC 85
Db	104 SKPCWYVYKARKFTSBCSVFVCS 127
RESULT 12	
TPA_MOUSE	
ID	TPA_MOUSE STANDARD; PRT; 559 AA.
AC	P11214; Q91VP2;
DT	01-JUL-1989 (Rel. 11, Created)
DT	15-SEP-2003 (Rel. 42, Last sequence update)
DT	15-SEP-2003 (Rel. 42, Last annotation update)
DE	Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE	(t-PA) (t-plasminogen activator).
GN	PlAT.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
PI	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=88087303; PubMed=2826484;
RA	Rickles R.J., Darrow A.L., Strickland S.;
RT	"Molecular cloning of complementary DNA to mouse tissue plasminogen
RT	activator mRNA and its expression during F9 teratocarcinoma cell
RT	differentiation.";
RL	J. Biol. Chem. 263:1563-1569 (1988).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Mammary gland;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McSwain P.J., McKernan K.J., Walek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko S., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -I- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZMOGEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -I- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -I- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -I- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -I- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-308 CATALYZED BY PLASMIN. TISSUE KALLIKREIN OR FACTOR XA.
CC -I- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: Contains 1 EGF-like domain.
CC -I- SIMILARITY: Contains 1 fibronectin type I domain.
CC -I- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J03520; AAA40470.1; --
DR EMBL; BC011256; AAH11256.1; --
DR PIR; A29941; A29941.
DR HSP; P00750; 1A5H.
DR MEROPS; S01-232; --
DR MGD; MGI:97610; Plat.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser-protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPsin_DOM; 1.
DR PROSITE; PS00134; TRYPsin_HIS; 1.
DR PROSITE; PS00135; TRYPsin_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
FT Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 17 PROBABLE.
FT PROPEP 18 29
FT CHAIN 30 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 30 308 TISSUE-TYPE PLASMINOGEN ACTIVATOR A
FT CHAIN 309 559 TISSUE-TYPE PLASMINOGEN ACTIVATOR B

FT DOMAIN 36 78 CHAIN
FT DOMAIN 79 117 FIBRONECTIN TYPE-1.
FT DOMAIN 124 205 EGF-LIKE.
FT DOMAIN 213 294 KRINGLE 1.
FT DOMAIN 309 355 SERINE PROTEASE.
FT ACT_SITE 355 355 CHARGE RELAY SYSTEM.
FT ACT_SITE 404 404 CHARGE RELAY SYSTEM.
FT ACT_SITE 510 510 CHARGE RELAY SYSTEM.
FT DISULFID 38 68 BY SIMILARITY.
FT DISULFID 66 75 BY SIMILARITY.
FT DISULFID 83 94 BY SIMILARITY.
FT DISULFID 88 105 BY SIMILARITY.
FT DISULFID 107 116 BY SIMILARITY.
FT DISULFID 124 205 BY SIMILARITY.
FT DISULFID 145 187 BY SIMILARITY.
FT DISULFID 176 200 BY SIMILARITY.
FT DISULFID 213 294 BY SIMILARITY.
FT DISULFID 234 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 297 428 INTERCHAIN (BY SIMILARITY).
FT DISULFID 340 356 BY SIMILARITY.
FT DISULFID 348 417 BY SIMILARITY.
FT DISULFID 442 516 BY SIMILARITY.
FT DISULFID 474 490 BY SIMILARITY.
FT DISULFID 506 534 BY SIMILARITY.
FT CARBOHYD 149 149 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 260 260 G -> A (IN REF. 1).
FT CONFLICT 325 325 P -> A (IN REF. 1).
SQ SEQUENCE 559 AA; 63122 MW; 8CCBE2BDB94514D9 CRC64;
Query Match 42.9%; Score 219; DB 1; Length 559;
Best Local Similarity 46.0%; Pred. No. 9.3e-18;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;
QY 2 TCYEGNGHFYRGKASTDTMGRCPLPWN SATVLOQTYHAHSDALQLGLGHKHYCRNPNR 61
DB 123 TCPEEQGITRYGTWSTAESGAECINWNSSVLSPKPNARPNALKLGHNHYCRNPNR 182
QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 183 LKPCWYVFKAGKYTFTEFCSTPACPKGK 209
RESULT 13
ID URT1 DESRO STANDARD; PRT; 477 AA.
AC P98119;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Salivary plasminogen activator alpha 1 precursor (EC 3.4.21.68) (DSPA
DE alpha-1).
OS Desmodus rotundus (Vampire bat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae;
OC Desmodontinae; Desmodus.
OX NCBI_TaxID=9430;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=92039036; PubMed=1937019;
RA Kraetzschmar J., Haendler B., Langer G., Boidol W., Bringmann P.,
RA Alagon A., Donner P., Schleuning W.D.;
RT "the plasminogen activator family from the salivary gland of the
RT vampire bat Desmodus rotundus: cloning and expression.";
RL Gene 105:229-237(1991).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=93393059; PubMed=1309059;
RA Schleuning W.-D., Alagon A., Boidol W., Bringmann P., Petri T.,
RA Kraetzschmar J., Haendler B., Langer G., Baldus B., Witt W.,

```

RA  Donner P.;
RT  "Plasminogen activators from the saliva of Desmodus rotundus (common
RL  vampire bat): unique fibrin specificity.";
RN  Ann. N.Y. Acad. Sci. 667:395-403(1992).
RC  [3]
RD  X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RE  TISSUE=salivary gland;
RF  MEDLINE=9802741; PubMed=9354616;
RG  Renatus M., Stubbs M.T., Huber R., Bringmann P., Donner P.,
RH  Schleuning W.D., Bode W.;
RI  "Catalytic domain structure of vampire bat plasminogen activator: a
RJ  molecular paradigm for proteolysis without activation cleavage.";
RK  Biochemistry 36:13483-13493(1997).
RL  -!- FUNCTION: PROBABLY ESSENTIAL TO SUPPORT THE FEEDING HABITS OF THIS
RM  EXCLUSIVELY HAEMATOPHAGOUS ANIMAL. POTENT THROMBOLYTIC AGENT.
RN  -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
RO  plasminogen to form plasmin.
RP  -!- ENZYME REGULATION: ACTIVATION OF ARG-|-VAL bond in
RQ  THE PRESENCE OF FIBRIN I.
RS  -!- SUBUNIT: Monomer.
RT  -!- DOMAIN: THE FIBRONECTIN TYPE-I DOMAIN MEDIATES BINDING TO FIBRIN,
RU  AND THE KRINGLE DOMAIN APPARENTLY MEDIATES FIBRIN-INDUCED
RV  STIMULATION OF ACTIVITY.
RW  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
RX  -!- SIMILARITY: Contains 1 EGF-like domain.
RY  -!- SIMILARITY: Contains 1 fibronectin type I domain.
RZ  -!- SIMILARITY: Contains 1 kringle domain.
RA  -----
RB  This SWISS-PROT entry is copyright. It is produced through a collaboration
RC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
RD  the European Bioinformatics Institute. There are no restrictions on its
RE  use by non-profit institutions as long as its content is in no way
RF  modified and this statement is not removed. Usage by and for commercial
RG  entities requires a license agreement (See http://www.isb-sib.ch/announce/
RH  or send an email to license@isb-sib.ch).
RI  -----
RJ  EMBL; M63987; AAA31591.1; -.
RK  EMBL; M63986; AAA31592.1; -.
RL  PIR; JS0597; JS0597.
RM  PDB; 1A51; 23-MAR-99.
RN  MEROPS; S01.232; -.
RO  Glycosylated; P98119; -.
RP  InterPro; IPR001314; Chymotrypsin.
RQ  InterPro; IPR006209; EGF like.
RS  InterPro; IPR000083; Fibronctnl.
RT  InterPro; IPR006210; IEGF.
RU  InterPro; IPR000001; Kringle.
RV  InterPro; IPR001254; Ser_protease_Try.
RW  Pfam; PF00008; EGF; 1.
RX  Pfam; PF00039; fn1; 1.
RY  Pfam; PF00051; kringle; 1.
RZ  Pfam; PF00089; trypsin; 1.
RA  PRINTS; PR00722; CHYMOTRYPSIN.
RB  PRINTS; PR00018; KRINGLE.
RC  ProDom; PD000395; Kringle; 1.
RD  SMART; SM00181; EGF; 1.
RE  SMART; SM00058; FNI; 1.
RF  SMART; SM00130; KR; 1.
RG  SMART; SM00020; Tryp_SPC; 1.
RH  PROSITE; PS00022; EGF_1; 1.
RI  PROSITE; PS01186; EGF_2; 1.
RJ  PROSITE; PS01253; FIBRONECTIN_1; 1.
RK  PROSITE; PS00021; KRINGLE_1; 1.
RL  PROSITE; PS50070; KRINGLE_2; 1.
RM  PROSITE; PS50240; TRYPSIN_DOM; 1.
RN  PROSITE; PS00134; TRYPSIN_HIS; 1.
RO  PROSITE; PS00135; TRYPSIN_SER; 1.
RP  Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
RQ  Kringle; EGF-like domain; Signal; Multigene family; 3D-structure.
RS  STGNAL 1 36 POTENTIAL.
RT  CHAIN 37 477 SALIVARY PLASMINOGEN ACTIVATOR ALPHA 1.
RU  DOMAIN 40 82 FIBRONECTIN TYPE-I.
RV  DOMAIN 83 121 EGF-LIKE.

```

```

FT  DOMAIN 128 209 KRINGLE.
FT  ACT_SITE 225 477 SERINE PROTEASE.
FT  ACT_SITE 272 477 CHARGE RELAY SYSTEM.
FT  ACT_SITE 321 477 CHARGE RELAY SYSTEM.
FT  ACT_SITE 428 477 CHARGE RELAY SYSTEM.
FT  DISULFID 42 72 BY SIMILARITY.
FT  DISULFID 70 79 BY SIMILARITY.
FT  DISULFID 87 98 BY SIMILARITY.
FT  DISULFID 92 109 BY SIMILARITY.
FT  DISULFID 111 120 BY SIMILARITY.
FT  DISULFID 128 209 BY SIMILARITY.
FT  DISULFID 149 191 BY SIMILARITY.
FT  DISULFID 180 204 BY SIMILARITY.
FT  DISULFID 214 345 BY SIMILARITY.
FT  DISULFID 257 273 BY SIMILARITY.
FT  DISULFID 265 334 BY SIMILARITY.
FT  DISULFID 359 434 BY SIMILARITY.
FT  DISULFID 391 407 BY SIMILARITY.
FT  DISULFID 424 452 BY SIMILARITY.
FT  CARBOHYD 153 153 N-LINKED (GLCNAC. .).
FT  CARBOHYD 398 398 /FTid=CAR_000027.
FT  CARBOHYD 398 398 N-LINKED (GLCNAC. .).
FT  CARBOHYD 398 398 /FTid=CAR_000028.
FT  TURN 214 215
FT  STRAND 223 224
FT  TURN 226 227
FT  STRAND 230 231
FT  HELIX 234 236
FT  TURN 238 239
FT  STRAND 240 245
FT  STRAND 254 263
FT  TURN 264 265
FT  STRAND 266 269
FT  HELIX 271 273
FT  TURN 280 282
FT  STRAND 284 287
FT  TURN 297 298
FT  STRAND 300 309
FT  TURN 311 312
FT  TURN 315 317
FT  TURN 319 320
FT  STRAND 323 328
FT  STRAND 338 338
FT  TURN 339 340
FT  STRAND 341 341
FT  STRAND 345 345
FT  TURN 349 350
FT  TURN 355 356
FT  STRAND 358 363
FT  STRAND 366 366
FT  STRAND 374 374
FT  STRAND 379 385
FT  HELIX 388 390
FT  TURN 393 398
FT  TURN 403 404
FT  STRAND 405 409
FT  TURN 425 426
FT  TURN 428 429
FT  STRAND 431 436
FT  TURN 437 438
FT  STRAND 439 448
FT  TURN 455 456
FT  STRAND 459 463
FT  HELIX 464 467
FT  HELIX 468 474
SQ  SEQUENCE 477 AA; 53616 MW; AA06FD1739C10B5E CRC64;
Query Match 41.8%; Score 213; DB 1; Length 477;
Best Local Similarity 46.4%; Pred. No. 3.9e-17;
Matches 39; Conservative 10; Mismatches 35; Indels 0; Gaps 0;
2 TCYEGNGHFYRGKASTDTMGRCPLPWNSATVLQOYTHAHRSDALQGLGKHNYCRNPQNR 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
: ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| :

```

```
Db 127 TCYEGQGVTVRGWTWSTAESRVECEINWNSSLLTRTYNGRMPDAPNLGLGNHNYCRNPNGA 186
QY 62 RRPWCYVQVGLKPLVQECMVHDC 85
Db 187 PKWCYVYKAGKFTSESCSPVCS 210

RESULT 14
ID TPA_BOVIN STANDARD; PRT; 566 AA.
AC Q28I98;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tissue-type plasminogen activator precursor (EC 3.4.21.68) (tPA)
DE (t-PA) (t-plasminogen activator) .
GN PLAT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA Ravn P., Berglund L., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen activators uPA
RL Int. Dairy J. 5:605-617(1995).."
CC -!- FUNCTION: CONVERTS THE ABUNDANT, BUT INACTIVE, ZYMOMEN PLASMINOGEN
CC TO PLASMIN BY HYDROLYZING A SINGLE ARG-VAL BOND IN PLASMINOGEN. BY
CC CONTROLLING PLASMIN-MEDIATED PROTEOLYSIS, IT PLAYS AN IMPORTANT
CC ROLE IN TISSUE REMODELING AND DEGRADATION, IN CELL MIGRATION AND
CC MANY OTHER PHYSIOLOGICAL EVENTS.
CC -!- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -!- SUBUNIT: HETERODIMER OF CHAIN A AND CHAIN B HELD BY A DISULFIDE
CC BOND.
CC -!- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR.
CC -!- PTM: THE SINGLE CHAIN, ALMOST FULLY ACTIVE ENZYME, CAN BE FURTHER
CC PROCESSED INTO A TWO-CHAIN FULLY ACTIVE FORM BY A CLEAVAGE AFTER
CC ARG-314 CATALYZED BY PLASMIN, TISSUE KALLIKREIN OR FACTOR XA.
CC -!- MISCELLANEOUS: BINDS TO THE KRINGLE STRUCTURE OF THE FIBRIN A
CC CHAIN. BINDING TO FIBRIN ENHANCES ITS CATALYTIC ACTIVITY.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
CC -!- SIMILARITY: Contains 2 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X85800; CAA59795.1; -.
DR HSPG; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibinctnl.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
```

```
DR SMART; SM00181; EGF; 1.
DR SMART; SM00058; FNI; 1.
DR SMART; SM00130; KR; 2.
DR SMART; SM00020; Tryp_spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01185; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 2.
DR PROSITE; PS50070; TRYPSIN_DOM; 1.
DR PROSITE; PS50240; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_SER; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolase; Serine protease; Glycoprotein;
KW Plasma; Kringle; EGF-like domain; Repeat; Signal.
FT SIGNAL 1 21
FT PROPEP 22 33
FT CHAIN 34 566
FT CHAIN 34 314
FT CHAIN 315 566
FT CHAIN 40 82
FT DOMAIN 83 121
FT DOMAIN 128 209
FT DOMAIN 219 300
FT DOMAIN 315 566
FT ACT_SITE 361 361
FT ACT_SITE 410 410
FT ACT_SITE 517 517
FT DISULFID 42 72
FT DISULFID 70 79
FT DISULFID 87 98
FT DISULFID 92 109
FT DISULFID 111 120
FT DISULFID 128 209
FT DISULFID 149 191
FT DISULFID 180 204
FT DISULFID 219 300
FT DISULFID 240 282
FT DISULFID 271 295
FT DISULFID 303 434
FT DISULFID 346 362
FT DISULFID 354 423
FT DISULFID 448 523
FT DISULFID 480 496
FT DISULFID 513 541
FT CARBOHYD 153 153
FT CARBOHYD 487 487
SQ SEQUENCE 566 AA; 63701 MW; 2EB6BEB4E32276C3 CRC64;

Query Match 41.0%; Score 209; DB 1; Length 566;
Best Local Similarity 47.6%; Pred. No. 1.3e-16;
Matches 39; Conservative 9; Mismatches 34; Indels 0; Gaps 0;

Oy 3 CYEGNGHFYRGKASTDTMGPRCLPWN SATV LQQYTHAHRSDALQGLGKHNYCRNPDRR 62
Db 219 CYTNGLAYRGTTRHTKSGASCLPWN SVFLTSKIYTAWSNAPALGLGKHNYCRNP 278

Oy 63 RPPCYVQVGLKPLVQECMVHDC 84
Db 279 QPWCHVWKDRQLTWYECVDPQC 300

RESULT 15
UROK_CHICK
ID UROK_CHICK STANDARD; PRT; 434 AA.
AC P15120;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Urokinase-type plasminogen activator precursor (EC 3.4.21.73) (uPA)
DE (U-plasminogen activator) .
DE Gallus gallus (Chicken).
```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP MEDLINE=90110185; PubMed=2295632;
RX Leslie N.D., Kessler C.A., Bell S.M., Degen J.L.;
RA "The chicken urokinase-type plasminogen activator gene.";
RL J. Biol. Chem. 265:1339-1344(1990).
CC -1- CATALYTIC ACTIVITY: Specific cleavage of Arg-|-Val bond in
CC plasminogen to form plasmin.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- SIMILARITY: Contains 1 EGF-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05187; AAA49131.1; -;
DR EMBL; J05188; AAA49130.1; -;
DR PIR; A35005; A35005.
DR HSP; P00763; IDPO.
DR MEROPS; S01.231; -;
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser.protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRODOM; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Plasminogen activation; Hydrolyase; Serine protease; Glycoprotein;
KW Kringle; EGF-like domain; Signal; Zymogen.
FT SIGNAL 1 20
FT CHAIN 21 434 UROKINASE-TYPE PLASMINOGEN ACTIVATOR.
FT CHAIN 21 171 CHAIN A (BY SIMILARITY).
FT CHAIN 173 434 CHAIN B (BY SIMILARITY).
FT DOMAIN 36 72 EGF-LIKE.
FT DOMAIN 79 158 KRINGLE.
FT DOMAIN 159 172 CONNECTING PEPTIDE.
FT DOMAIN 173 434 SERINE PROTEASE.
FT DISULFID 40 48 BY SIMILARITY.
FT DISULFID 42 60 BY SIMILARITY.
FT DISULFID 62 71 BY SIMILARITY.
FT DISULFID 162 296 INTERCHAIN (BY SIMILARITY).
FT DISULFID 202 218 BY SIMILARITY.
FT DISULFID 210 285 BY SIMILARITY.
FT DISULFID 310 379 BY SIMILARITY.
FT DISULFID 342 358 BY SIMILARITY.
FT DISULFID 369 397 BY SIMILARITY.
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 272 272 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 373 373 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 228 228 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 434 AA; 49400 MW; BD861048DD666A55 CRC64;

Query Match 38.9%; Score 198.5; DB 1; Length 434;
Best Local Similarity 54.4%; Pred. No. 1.7e-15;
Matches 37; Conservative 7; Mismatches 19; Indels 5; Gaps 2;
QY 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLQ-QTYHAHRSDALQLGLGKHNYCRNPDR 61
DB 79 CYSNGGDIYRGMAEDP---GCLYWDHPSVIRWGDYHADLKNALQLGLGKHNYCRNPGR 134
QY 62 RRPWCYVQ 69
DB 135 SRPWCYTK 142
RESULT 16
HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
DE activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Liver, and Serum;
RX MEDLINE=93252878; PubMed=7683665;
RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,
RA Kitamura N.;
RT "Molecular cloning and sequence analysis of the cDNA for a human
RT serine protease responsible for activation of hepatocyte growth
RT factor. Structural similarity of the protease precursor to blood
RT coagulation factor XII.";
RL J. Biol. Chem. 268:10024-10028(1993).
RN [2]
RP SEQUENCE OF 40-655 FROM N.A.
RA Zhao S., Odell C.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ACTIVATES HEPATOCYTE GROWTH FACTOR (HGF) BY
CC CONVERTING IT FROM A SINGLE CHAIN TO A HETERODIMERIC FORM.
CC -1- SUBUNIT: DIMER OF A SHORT CHAIN AND A LONG CHAIN LINKED BY A
CC DISULFIDE BOND.
CC -1- SUBCELLULAR LOCATION: SECRETED AS AN INACTIVE SINGLE-CHAIN
CC PRECURSOR AND IS THEN ACTIVATED TO A HETERODIMERIC FORM.
CC -1- TISSUE SPECIFICITY: LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: Contains 2 EGF-like domains.
CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
CC -1- SIMILARITY: Contains 1 kringle domain.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D14012; BAA03113.1; -;
DR EMBL; Z69923; CAA93803.1; -;
DR PIR; A46688; A46688.
DR HSP; P00763; IDPO.
DR MEROPS; S01.228; -;
DR Genew; HGNC:4894; HGFA.
DR MIM; 604552; -;
DR GO; GO:0005576; C:extracellular; TAS.

DR InterPro; IPR000083; Fibrinctnl.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR008210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 DR PRINTS; PRO0013; FNTYPEII.
 DR PRINTS; PRO0018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS00070; KRINGLE_2; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;
 KW EGF-like domain; Repeat; Zymogen.
 FT SIGNAL 1 29 BY SIMILARITY.
 FT PROPEP 30 369 CLEAVED IN ACTIVE FORM (BY SIMILARITY).
 FT CHAIN 370 405 HEPATOCYTE GROWTH FACTOR ACTIVATOR SHORT CHAIN.
 FT CHAIN 406 653 HEPATOCYTE GROWTH FACTOR ACTIVATOR LONG CHAIN.
 FT DOMAIN 105 145 FIBRONECTIN TYPE-II.
 FT DOMAIN 157 195 EGF-LIKE 1.
 FT DOMAIN 197 237 FIBRONECTIN TYPE-I.
 FT DOMAIN 238 276 EGF-LIKE 2.
 FT DOMAIN 283 364 KRINGLE.
 FT DOMAIN 406 653 SERINE PROTEASE.
 FT ACT_SITE 445 445 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 495 495 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 596 596 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 105 130 BY SIMILARITY.
 FT DISULFID 119 145 BY SIMILARITY.
 FT DISULFID 161 172 BY SIMILARITY.
 FT DISULFID 166 183 BY SIMILARITY.
 FT DISULFID 185 194 BY SIMILARITY.
 FT DISULFID 199 227 BY SIMILARITY.
 FT DISULFID 225 234 BY SIMILARITY.
 FT DISULFID 242 253 BY SIMILARITY.
 FT DISULFID 247 264 BY SIMILARITY.
 FT DISULFID 266 275 BY SIMILARITY.
 FT DISULFID 283 364 BY SIMILARITY.
 FT DISULFID 304 346 BY SIMILARITY.
 FT DISULFID 335 359 BY SIMILARITY.
 FT DISULFID 392 519 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 430 446 BY SIMILARITY.
 FT DISULFID 438 508 BY SIMILARITY.
 FT DISULFID 533 602 BY SIMILARITY.
 FT DISULFID 565 581 BY SIMILARITY.
 FT DISULFID 592 620 BY SIMILARITY.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 544 544 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 164 164 G -> W (IN REF. 2).
 SQ SEQUENCE 653 AA; 70567 MW; 38B4B20255DF7FDC CRC64;

Query Match 36.9%; Score 188; DB 1; Length 653;
 Best Local Similarity 53.0%; Pred No. 4.2e-14;
 Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;
 Qy 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLTQYTHAHRSDALQLGLGKHNCRPNRR 62
 Db 283 CFLGNGTEYRGVASTAASGLSLAUNSDLLVQLHVDVAAAVLLGLGPHAYCRNPKDE 342
 Qy 63 RWCYV 68
 Db 343 RWCYV 348
 RESULT 18
 FA12_CAVPO STANDARD; PRT; 603 AA.
 ID FA12_CAVPO
 AC Q04962;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
 DE (HAF) (Fragment).
 GN F12.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
 RC TISSUE=Liver;
 RX MEDLINE=93003367; PubMed=1390917;
 RA Semba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
 RA Kambara T., Okabe H.;
 RT "Primary structure of guinea-pig Hageman factor: sequence around the
 RT cleavage site differs from the human molecule.";
 RL Biochim. Biophys. Acta 1159:113-121(1992).
 CC -!- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
 CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-Ile bonds in factor
 CC VII to form factor VIIa and factor XI to form factor Xla.
 CC -!- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -!- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -!- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -!- SIMILARITY: Contains 2 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X68615; CAA48600.1; -;
 CC PIR; S28941; S28941.
 CC HSP; P00763; IDPO.
 CC MEROPS; S01.211; -;
 CC InterPro; IPR001314; Chymotrypsin.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR006209; EGF_like.
 CC InterPro; IPR000083; Fibrinctnl.
 CC InterPro; IPR000562; FN_type_II.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR000001; Kringle.
 CC InterPro; IPR001254; Ser_protease_Try.


```
RESULT 20
APOA APOA MACMU STANDARD; PRT; 1420 AA.
AC P1417;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Apolipoprotein(a) (EC 3.4.21.-) (Apo(a)) (Lp(a)) (Fragment).
GN Lp(a).
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=89174660; PubMed=2925643;
RA Tomlinson J.E., McLean J.W., Lawn R.M.;
RT "Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of
RT synthesis.";
RL J. Biol. Chem. 264:5957-5965 (1989).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibronectin and
CC decorin (by similarity).
CC -!- PTM: N- and O-glycosylated (By similarity).
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrinogen binding.
CC -!- MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation (By similarity)
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains at least 10 kringle domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; J04635; AAA36833.1; -
DR PIR; A32869; A32869.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.226; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 11.
DR Pfam; PF00083; trypsin_1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 10.
DR SMART; SM00130; KR; 10.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 10.
DR PROSITE; PS50070; KRINGLE_2; 10.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
DR PROSITE; PS00135; TRYPSIN_SER; FALSE NEG.
DR Hydrolase; Serine protease; Lipid transport; Plasma; Glycoprotein;
KW Kringle; Repeat; Atherosclerosis.
FT NON_TER 1
FT DOMAIN 49 127 KRINGLE 1.
FT DOMAIN 163 241 KRINGLE 2.
```

```
FT DOMAIN 277 355 KRINGLE 3.
FT DOMAIN 391 469 KRINGLE 4.
FT DOMAIN 505 583 KRINGLE 5.
FT DOMAIN 619 697 KRINGLE 6.
FT DOMAIN 725 803 KRINGLE 7.
FT DOMAIN 839 917 KRINGLE 8.
FT DOMAIN 953 1031 KRINGLE 9.
FT DOMAIN 1067 1145 KRINGLE 10.
FT DOMAIN 1191 1420 SERINE PROTEASE.
SQ SEQUENCE 1420 AA; 158367 MW; BE102949E03CB0E CRC64;

Query Match 31.6%; Score 161; DB 1; Length 1420;
Best Local Similarity 40.9%; Pred. No. 1.3e-10;
Matches 36; Conservative 9; Mismatches 31; Indels 12; Gaps 4;

QY 3 CYEGNGHFYRGKASTDTWGRPCLPWSATVLQ--QTYHAHRS DALQLGLGKHNYCRNPDN 60
DB 1068 CYHGNGSYRGFTSTVTGTCQSSSWTPHQHKRTPEHPNDLLTM-----NYCRNPD 1122
QY 61 RRRPWCYVQVGLKPLVQE--CMVHDCAD 86
DB 1123 DTGFWCFT--MDPSVRREYCNLTRCSD 1147

RESULT 21
KRM1 MOUSE STANDARD; PRT; 473 AA.
ID KRM1 MOUSE
AC Q99N43;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A., DEVELOPMENTAL STAGE, AND TISSUE SPECIFICITY.
TISSUE=Brain, and Kidney;
MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Takahashi T., Matsumoto K.,
RA Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel
RT kringle-containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: In the adult, widely expressed with high
CC levels in heart, lung, kidney, skeletal muscle and testis.
CC -!- DEVELOPMENTAL STAGE: In the embryo, expression is first detected
CC on day 9 and increases up to day 18. Lower levels are found in
CC ectodermal ridge (AER) of the developing fore- and hindlimb buds,
CC the telencephalon and the first brachial arch. At 10.5 dpc,
CC expression is also observed in the myotome and in sensory tissues
CC such as the nasal pit and optic vesicle.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AB059617; BAB40968.1; -
DR HSP; P00747; ICEA.
```

```

DR MGD; MGI:1933988; Kremen.
DR GO; GO:0016021; C: integral to membrane; NAS.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00021; KRINGLE_2; 1.
DR PROSITE; PS00070; Kringle.
DR PROSITE; PS00070; Kringle; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 473 KREMEN PROTEIN 1.
FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 393 413 POTENTIAL.
FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 114 KRINGLE.
FT DOMAIN 120 210 WSC.
FT DOMAIN 214 321 CUB.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 473 AA; 51716 MW; 586827788BE3FDD1 CRC64;

Query Match 31.2%; Score 159; DB 1; Length 473;
Best Local Similarity 45.6%; Pred. No. 6.7e-11;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

QY 3 CYEGNGHYRGKASTDTM--GRPLPWNSTVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60
DB 32 CFTANGADYRGTSWTALQGGKPCLFWNE--TFQHPYNTLKYPNEGGLGHEHNYCRNPDG 89

QY 61 RRRPWCYV 68
DB 90 DVSPWCYV 97

RESULT 22
KRM1_RAT
ID KRM1_RAT STANDARD; PRT; 473 AA.
AC Q924S4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1 OR KREMEN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.; Nakamura T.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; AB065090; BAB62003.1; -.
CC GO; GO:0016021; C: integral to membrane; ISS.
CC InterPro; IPR000859; CUB_domain.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE_1; 1.
CC PROSITE; PS00021; KRINGLE_2; 1.
CC PROSITE; PS00070; Kringle.
CC PROSITE; PS00070; Kringle; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 473 KREMEN PROTEIN 1.
FT DOMAIN 21 392 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 393 413 POTENTIAL.
FT DOMAIN 414 473 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 114 KRINGLE.
FT DOMAIN 120 210 WSC.
FT DOMAIN 214 321 CUB.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 255 255 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 473 AA; 51869 MW; 9B510857DF856F08 CRC64;

Query Match 31.2%; Score 159; DB 1; Length 473;
Best Local Similarity 45.6%; Pred. No. 6.7e-11;
Matches 31; Conservative 7; Mismatches 26; Indels 4; Gaps 2;

QY 3 CYEGNGHYRGKASTDTM--GRPLPWNSTVLQQTTHAHRSDALQLGLGKHNYCRNPDN 60
DB 32 CFTANGADYRGTSWTALQGGKPCLFWNE--TFQHPYNTLKYPNEGGLGHEHNYCRNPDG 89

QY 61 RRRPWCYV 68
DB 90 DVSPWCYV 97

RESULT 23
KRM1_XENLA
ID KRM1_XENLA STANDARD; PRT; 452 AA.
AC Q90Y90;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
DE and the nose) (Dickkopf receptor).
GN KREMEN1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Nakamura T.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Contains 1 CUB domain.

```

```

CC CC -!- SIMILARITY: Contains 1 kringle domain.
CC CC -!- SIMILARITY: Contains 1 WSC domain.
CC CC -----
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AB070851; BAB64294.1; -
CC CC InterPro; IPR000859; CUB domain.
CC CC InterPro; IPR000001; Kringle.
CC CC InterPro; IPR002889; WSC.
CC CC Pfam; PF00431; CUB; 1.
CC CC Pfam; PF00051; kringle; 1.
CC CC Pfam; PF01822; WSC; 1.
CC CC PRINTS; PR00018; KRINGLE.
CC CC ProDom; PD000395; Kringle; 1.
CC CC SMART; SM00042; CUB; 1.
CC CC SMART; SM00130; KR; 1.
CC CC SMART; SM00321; WSC; 1.
CC CC PROSITE; PS01180; CUB; 1.
CC CC PROSITE; PS00021; KRINGLE_1; 1.
CC CC PROSITE; PS00070; KRINGLE_2; 1.
CC CC Wnt signaling pathway; Glycoprotein; Kringle;
CC CC Transmembrane; Signal.
CC CC SIGNAL 1 22
CC CC CHAIN 23 452
CC CC DOMAIN 23 369
CC CC TRANSMEM 370 390
CC CC DOMAIN 391 452
CC CC DOMAIN 29 112
CC CC DOMAIN 118 208
CC CC DOMAIN 212 319
CC CC CARBOHYD 43 43
CC CC CARBOHYD 57 57
CC CC CARBOHYD 215 215
CC CC CARBOHYD 253 253
CC CC CARBOHYD 291 291
CC CC CARBOHYD 328 328
CC CC CARBOHYD 344 344
CC CC SEQUENCE 452 AA; 50188 MW; ED24BCD1AF4564E2 CRC64;
CC CC
CC CC Query Match 31.0%; Score 158; DB 1; Length 452;
CC CC Best Local Similarity 44.1%; Pred. No. 8.4e-11;
CC CC Matches 30; Conservative 9; Mismatches 25; Indels 4; Gaps 2;
CC CC
CC CC Qy 3 CYEGNGHYRGKASTDTM--GRPCLPMSATVLQOTYHAHSDALQLGLGKHNYCRNPDN 60
CC CC Db 30 CYTVNGADYRTQNTSLDGGKPCLFWE--TFQHPYNTLKYPNGGGLGKHNYCRNPDG 87
CC CC Qy 61 RRRPWCYV 68
CC CC Db 88 DVSPWCYI 95
CC CC
CC CC RESULT 24
CC CC KRM1_HUMAN
CC CC ID KRM1_HUMAN STANDARD; PRT: 475 AA.
CC CC AC Q96M08; Q9BY70; Q9UGS5; Q9UGU1;
CC CC DT 28-FEB-2003 (Rel. 41, Created)
CC CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC CC DE Kremen protein 1 precursor (Kringle-containing protein marking the eye
CC CC and the nose) (Dickkopf receptor).
CC CC GN KREMEN1 OR KREMEN.
CC CC OS Homo sapiens (Human).
CC CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC CC OX NCBI_TaxID=9606;
CC CC RN [1]

```

```

RP RP SEQUENCE FROM N.A. (ISOFORM 1).
RA RA Nakamura T., Nakamura T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN RN [2]
RP RP SEQUENCE FROM N.A. (ISOFORM 2).
RA RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Sekine K., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Nishikawa T., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kaneshori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN RN [3]
RP RP SEQUENCE FROM N.A.
RX RX MEDLINE=20057165; PubMed=10591208;
RA RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Courty D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.P., Levensha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon I., Steward C.A., Sulston J.F., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.H., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA Emanuel H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tikhunov Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
CC CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC CC to block Wnt/beta-catenin signaling (By similarity).
CC CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC CC -!- ALTERNATIVE PRODUCTS:
CC CC Event=Alternative splicing; Named isoforms=2;
CC CC Name=1;
CC CC IsoId=Q96M08-1; Sequence=Displayed;
CC CC Name=2;
CC CC IsoId=Q96M08-2; Sequence=VSP_003900;
CC CC Note=No experimental confirmation available;
CC CC -!- SIMILARITY: Contains 1 CUB domain.
CC CC -!- SIMILARITY: Contains 1 kringle domain.

```

```

CC CC -!- SIMILARITY: Contains 1 WSC domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL; AB059618; BAB40969.1; -
DR EMBL; AK056425; BAB71180.1; -
DR EMBL; Z95116; CAB62952.1; -
DR EMBL; AL021393; CAB62959.1; -
DR GenBank; HGNC:17550; KREMEN1.
DR GO; GO:0016021; C: integral to membrane; ISS.
DR GO; GO:0005624; C: membrane fraction; TAS.
DR GO; GO:0007154; P: cell communication; TAS.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS00070; KRINGLE 2; 1.
KW Wnt signaling pathway; Signal; Transmembrane; Kringle;
KW Alternative splicing.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 475 KREMEN PROTEIN 1.
FT DOMAIN 21 394 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 395 415 POTENTIAL.
FT DOMAIN 416 475 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 33 116 KRINGLE.
FT DOMAIN 121 202 WSC.
FT DOMAIN 216 323 CUB.
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 473 475 VSD -> AIQDSVTSLSWSQGPQPSI (in isoform 2).
FT FTID=VSP 003900.
FT CONFLICT 29 30 MISSING (IN REF. 1).
FT CONFLICT 206 206 I -> V (IN REF. 2).
SQ SEQUENCE 475 AA; 51898 MW; B7E86FD80F96A0A4 CRC64;
Query Match 30.6%; Score 156; DB 1; Length 475;
Best Local Similarity 44.1%; Pred. No. 1.5e-10;
Matches 30; Conservative 8; Mismatches 26; Indels 4; Gaps 2;
QY 3 CYEGNGHFYRGKASTDTM--GRPCLPWSATVLQQTVAHRSDAQLGLGKKNYCRNPND 60
Db 34 CFTANGADFRGTQNTWALQGRKPLFWNE--TFQHPYTLKYPNGEGGLGKKNYCRNPND 91
QY 61 RRPFCVY 68
Db 92 DVSPFCVY 99
RESULT 25
APOA HUMAN
ID APOA HUMAN STANDARD; PRT; 4548 AA.
AC P08519;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)

```

```

DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Apolipoprotein(a) precursor (EC 3.4.21.-) (Apo(a)) (Lp(a)).
GN LpA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88039109; PubMed=3670400;
RA McLean J.W., Tomlison J.E., Kuang W.-J., Eaton D.L., Chen E.Y.,
RA Fless G.W., Scannu A.M., Lawn R.M.;
RT "cDNA sequence of human apolipoprotein(a) is homologous to
RT plasminogen.";
RL Nature 330:132-137(1987).
RN [2]
RP SERINE PROTEASE ACTIVITY.
RX MEDLINE=90076123; PubMed=2531657;
RA Salonen E.-M., Jauhainen M., Zardi L., Vaheri A., Ehnholm C.;
RT "Lipoprotein(a) binds to fibrinogen and has serine proteinase
RT activity capable of cleaving it.";
RL EMBO J. 8:4035-4040(1989).
RN [3]
RP REVIEW.
RX MEDLINE=90049223; PubMed=2530631;
RA Utermann G.;
RT "The mysteries of lipoprotein(a).";
RL Science 246:904-910(1989).
RN [4]
RP CHARACTERIZATION OF THE N- AND O-LINKED GLYCANS.
RX MEDLINE=21303595; PubMed=11294842;
RA Garner B., Merry A.H., Royle L., Harvey D.J., Rudd P.M., Thillet J.;
RT "Structural elucidation of the N- and O-glycans of human
RT apolipoprotein(a): role of o-glycans in conferring protease
RT resistance.";
RL J. Biol. Chem. 276:22200-22208(2001).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 4121-4208.
RX MEDLINE=96217891; PubMed=8642595;
RA Mikol V., Lograsso P.V., Boettcher B.R.;
RT "Crystal structures of apolipoprotein(a) kringle IV37 free and
RT complexed with 6-aminohexanoic acid and with p-aminomethylbenzoic
RT acid: existence of novel and expected binding modes.";
RL J. Mol. Biol. 256:751-761(1996).
RN [6]
RP VARIANT ARG-4193.
RX MEDLINE=95002201; PubMed=7918682;
RA Scannu A.M., Pfaffinger D., Lee J.C., Hinman J.;
RT "A single point mutation (Trp72-->Arg) in human apo(a) kringle 4-37
RT associated with a lysine binding defect in Lp(a).";
RL Biochim. Biophys. Acta 1227:41-45(1994).
CC -!- FUNCTION: Apo(a) is the main constituent of lipoprotein(a)
CC (Lp(a)). It has serine proteinase activity and is able of
CC autolysis. Inhibits tissue-type plasminogen activator 1.
CC Lp(a) may be a ligand for megalin/Gp 330.
CC -!- SUBUNIT: Disulfide-linked to apo-B100. Binds to fibrinogen and
CC decorin.
CC -!- PTM: N- and O-glycosylated. The N-glycans are complex biantennary
CC structures present in either a mono- or disialylated state. The
CC O-glycans are mostly (80%) represented by the monosialylated core
CC type I structure, NeuNAcHex2-3GalHex1-3GalNAc, with smaller
CC amounts of disialylated and non-sialylated O-glycans also
CC detected.
CC -!- DISEASE: Elevated plasma concentrations of apo(a) and its
CC naturally occurring proteolytic fragments are correlated with
CC atherosclerosis. Homology with plasminogen kringle IV and V is
CC thought to underlie the atherogenicity of the protein, because the
CC fragments are competing with plasminogen for fibrinogen binding.
CC MISCELLANEOUS: Apo(a) is known to be proteolytically cleaved,
CC leading to the formation of the so called mini-Lp(a). Apo(a)
CC fragments accumulate in atherosclerotic lesions, where they may
CC promote thrombogenesis. O-glycosylation may limit the extent of
CC proteolytic fragmentation.

```


CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 EGF-like domains. S1.
 CC -1- SIMILARITY: Contains 1 fibronectin type I domain.
 CC -1- SIMILARITY: Contains 1 fibronectin type II domain.
 CC -1- SIMILARITY: Contains 1 kringle domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S70164; ABE30804.2; -;
 DR PIR; S45281; S45281.
 DR HSSP; P00763; 1DPO.
 DR MEROPS; S01.211; -;
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR006209; EGF-like.
 DR InterPro; IPR000083; Fibinact1.
 DR InterPro; IPR000562; FN_Type_II.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR000001; Kringle.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00039; fn1; 1.
 DR Pfam; PF00040; fn2; 1.
 DR Pfam; PF00051; kringle; 1.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00013; FNTYPEII.
 DR PRINTS; PR00018; KRINGLE.
 DR ProDom; PD000995; FN_Type_II; 1.
 DR ProDom; PD000395; Kringle; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00058; FN1; 1.
 DR SMART; SM00059; FN2; 1.
 DR SMART; SM00130; KR; 1.
 DR SMART; SM00020; Tryp_SPC; 1.
 DR PROSITE; PS00022; EGF_1; 2.
 DR PROSITE; PS01186; EGF_2; FALSE NEG.
 DR PROSITE; PS01253; FIBRONECTIN_1; 1.
 DR PROSITE; PS00023; FIBRONECTIN_2; 1.
 DR PROSITE; PS00021; KRINGLE_1; 1.
 DR PROSITE; PS50070; KRINGLE_2; 1.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
 KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
 FT NON_TER 1 1
 FT SIGNAL <1 9
 FT CHAIN 10 349
 FT CHAIN 350 593
 FT DOMAIN 37 78
 FT DOMAIN 84 121
 FT DOMAIN 123 163
 FT DOMAIN 164 200
 FT DOMAIN 207 287
 FT DOMAIN 297 333
 FT DOMAIN 350 593
 FT ACT_SITE 389 389
 FT ACT_SITE 438 438
 FT ACT_SITE 541 541
 FT DISULFID 88 100
 FT DISULFID 111 120
 FT DISULFID 125 153
 FT DISULFID 151 160
 FT DISULFID 168 179
 FT DISULFID 173 188
 FT DISULFID 190 199

FT DISULFID 207 287 BY SIMILARITY.
 FT DISULFID 230 269 BY SIMILARITY.
 FT DISULFID 258 282 BY SIMILARITY.
 FT DISULFID 336 463 BY SIMILARITY.
 FT DISULFID 374 390 BY SIMILARITY.
 FT DISULFID 382 452 BY SIMILARITY.
 FT DISULFID 413 416 BY SIMILARITY.
 FT DISULFID 479 547 BY SIMILARITY.
 FT DISULFID 510 526 BY SIMILARITY.
 FT DISULFID 537 568 BY SIMILARITY.
 FT CARBOHYD 99 99 O-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;
 Query Match 30.5%; Score 155.5; DB 1; Length 593;
 Best Local Similarity 38.4%; Pred. No. 2.2e-10;
 Matches 33; Conservative 9; Mismatches 37; Indels 7; Gaps 3;
 Qy 2 TCYE--GNGHFYRGKASTDTNGRCLPWN SATVLQOQY-HAHRSDALQLGKHNKCNRP 58
 Db 206 SCYDDRRGLSYRGMA GTTLGAPCQSWAS-----EATVWNVTAQVLNWLGLGDHAFRCNP 261
 Qy 59 DNRERPWCYQVGLKPLVQECWHDC 84
 Db 262 DNDTRPWCFFIWKGDRLSNWNYCLAPC 287
 RESULT 27
 PLMN RAT
 ID PLMN RAT STANDARD; PRT; 169 AA.
 AC Q01177;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Plasminogen (EC 3.4.21.7) (Fragment).
 GN PLG.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=91250378; PubMed=1645711;
 RA Kanalas J.J., Makker S.P.;
 RT "Identification of the rat Heymann nephritis autoantigen (GP330) as a
 RT receptor site for plasminogen."
 RL J. Biol. Chem. 266:10825-10829(1991).
 CC -1- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
 CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
 CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
 CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
 CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
 CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
 CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
 CC LAMININ AND VON WILLEBRAND FACTOR.
 CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
 CC higher selectivity than trypsin. Converts fibrin into soluble
 CC products.
 CC -1- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
 CC ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
 CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
 CC -1- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
 CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
 CC -1- SIMILARITY: Contains 5 kringle domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
```

```

DR DR EMBL; M62832; AAA41884.1; -.
DR DR PIR; A40522; A40522.
DR DR HSP; P00747; LPNK.
DR DR MEROPS; S01.233; -.
DR DR InterPro; IPR000001; Kringle.
DR DR InterPro; IPR003966; Prothrombin.
DR DR Pfam; PF00051; Kringle_2.
DR DR PRINTS; PR00018; KRINGLE.
DR DR PRINTS; PR01505; PROTHROMBIN.
DR DR ProDom; PD000395; Kringle; 2.
DR DR SMART; SM00130; KR; 1.
DR DR PROSITE; PS00021; KRINGLE 1; 1.
DR DR PROSITE; PS00070; KRINGLE 2; 2.
DR DR PROSITE; PS00240; TRYPSIN_DOM; PARTIAL.
DR DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Repeat.
FT NON_TER 1
FT DOMAIN <1 10 KRINGLE 3 (BY SIMILARITY).
FT DOMAIN 34 112 KRINGLE 4 (BY SIMILARITY).
FT DOMAIN 139 >169 KRINGLE 5 (BY SIMILARITY).
FT DISULFID 34 112 BY SIMILARITY.
FT DISULFID 55 95 BY SIMILARITY.
FT DISULFID 83 107 BY SIMILARITY.
FT NON_TER 169 169
SQ SEQUENCE 169 AA; 18401 MW; 77A54214C49D010C CRC64;

Query Match 29.3%; Score 149.5; DB 1; Length 169;
Best Local Similarity 37.8%; Pred. No. 2.9e-10;
Matches 34; Conservative 12; Mismatches 29; Indels 15; Gaps 6;

QY 3 CYEGNGHFRGKASTMTGRPCLPWNSATVLQOYTHAHRSDALQL---GLGKHNYCRNPD 59
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 34 CYQGNGKSVRGTSSTTTGKQSW-----VSNTPHSHKSTPANFPDSGL-EMNYCRNPD 97
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 60 N-RRRWYVQVGLKPLV--QECMVHDCAD 86
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 88 NDORGWPCFT---TDFSVRWYCNLKRCS 114
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 28
ID PLMN BOVIN STANDARD; PRT; 812 AA.
AC P06858; Q28162;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plasminogen precursor (EC 3.4.21.7).
GN PLG.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Berglund L., Andersen M.D., Petersen T.E.;
RT "Cloning and characterization of the bovine plasminogen cDNA.";
RL Int. Dairy J. 5:593-603 (1995).
RN [2]
RP SEQUENCE OF 27-812, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=85203906; PubMed=3846532;
RA Schaller J., Moser P.W., Danneberger-Muller G.A.K., Rosselet S.J.,
RA Kampfer U., Rickli E.E.;
RT "Complete amino acid sequence of bovine plasminogen. Comparison with
human plasminogen.";
```

```

RL RN Eur. J. Biochem. 149:267-278 (1985).
RP [3].
RX SEQUENCE OF 706-812 FROM N.A.
RX MEDLINE=85023311; PubMed=6148961;
RA Malinowski D.P., Sadler J.E., Davie E.W.;
RT "Characterization of a complementary deoxyribonucleic acid coding for
human and bovine plasminogen.";
RL Biochemistry 23:4243-4250 (1984).
RN [4].
RP CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=88185329; PubMed=3356193;
RA Marti T., Schaller J., Rickli E.E., Schmid K., Kamerling J.P.,
RA Gerwig G.J., van Halbeek H., Vliegenthart J.F.;
RT "The N- and O-linked carbohydrate chains of human, bovine and porcine
plasminogen. Species specificity in relation to sialylation and
fucosylation patterns.";
RL Eur. J. Biochem. 173:57-63 (1988).
CC -!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS
CC A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING
CC EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION,
CC AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE
CC GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN
CC ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH
CC AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN,
CC LAMININ AND VON WILLEBRAND FACTOR.
CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa;
CC higher selectivity than trypsin. Converts fibrin into soluble
CC products.
CC -!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN
CC ACTIVATORS. BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO
CC FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.
CC -!- PTM: N-LINKED GLYCAN CONTAIN N-ACETYLGLUCOSAMINE AND SIALIC ACID.
CC O-LINKED GLYCANS CONSIST OF GAL-GALNAc DISACCHARIDE WITH IS
CC MODIFIED WITH UP TO 2 SIALIC ACID RESIDUES (MICROHETEROGENEITY).
CC -!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN
CC IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.
CC -!- SIMILARITY: Contains 5 kringle domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

DR EMBL; X79402; CAA55939.1; -.
DR EMBL; K02935; AAA30714.1; -.
DR PIR; S45046; PLBO.
DR HSP; P00747; 2PK4.
DR MEROPS; S01.233; -.
DR GlycoSuiteDB; P06869; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; Kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 5.
DR SMART; SM00473; PAN AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE 1; 5.
DR PROSITE; PS00070; KRINGLE 2; 5.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
```

DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis;
KW Tissue remodeling; Blood coagulation; Kringle; Zymogen; Repeat;
KW SIGNAL
FT FT CHAIN 1 26
FT FT CHAIN 27 812
FT FT CHAIN 28 583
FT FT CHAIN 584 812
FT FT CHAIN 110 188
FT FT DOMAIN 192 269
FT FT DOMAIN 282 359
FT FT DOMAIN 384 461
FT FT DOMAIN 485 564
FT FT DOMAIN 584 812
FT FT CARBOHYD 315
FT FT CARBOHYD 365
FT FT ACT SITE 624
FT FT ACT SITE 667
FT FT ACT SITE 762
FT FT ACT SITE 335
FT FT CONFLICT 316
FT FT CONFLICT 516
FT FT CONFLICT 555
FT FT CONFLICT 744
FT FT CONFLICT 744
SQ SEQUENCE 812 AA; 91216 MW; 38AA6A691E220946 CRC64;
Query Match 29.2%; Score 149; DB 1; Length 812;
Best Local Similarity 39.3%; Pred. No. 1.7e-09;
Matches 35; Conservative 8; Mismatches 33; Indels 14; Gaps 5;
QY 3 CYEGNGHYRGKASTDTGRCPLPNSATVLTQITVHAH---RSDALQLGKHNKNCNPD 59
Db 384 CVHNGQSGYRGTSSTITGRKQSWSS-----MTPHRLKTPENYNAGL-TMNYCRNPD 437
QY 60 NRRRWPCYVQGLKPLV--QECWHDCCAD 86
Db 438 ADKSPWCYT---TDFRVRWFCNLKCKSE 463
RESULT 29
ID KRM2 HUMAN STANDARD; PRT: 462 AA.
AC Q8NCW0; Q8N2J4; Q8NCW1; Q96GL8; Q9BTP9;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Kremen protein 2 precursor (Kringle-containing protein marking the eye and the nose) (Dickkopf receptor 2).
GN KREMEN2 OR KRM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RA Tanaka S., Sugimachi K.;
RT "Human Kremen2 and Wnt signaling.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RA Tissue-Ovarian carcinoma;
RC Tissue-Ovarian carcinoma;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamanoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahara K.,
RA Masuhito Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RT "NED0 human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 4).
RC Tissue-Brain, and uterus;
RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrincci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marita M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Receptor for Dickkopf protein. Cooperates with Dickkopf
CC to block Wnt/beta-catenin signaling. Forms a ternary complex with
CC Dkk1 and LRP6 and induces rapid endocytosis and removal of the Wnt
CC receptor LRP6 from the plasma membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=Q8NCW0-1; Sequence=Displayed;
CC Name=2; Synonyms=Kremen2a;
CC IsoId=Q8NCW0-2; Sequence=VSP_050509, VSP_050510;
CC Name=3; Synonyms=Kremen2b;
CC IsoId=Q8NCW0-3; Sequence=VSP_050511, VSP_050512;
CC Name=4; Synonyms=Kremen2c;
CC IsoId=Q8NCW0-4; Sequence=VSP_050513, VSP_050514;
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: Contains 1 kringle domain.
CC -!- SIMILARITY: Contains 1 WSC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AB086405; BAC00872.1; -
CC EMBL; AB086355; BAC00823.1; -
CC EMBL; AB086356; BAC00824.1; -
CC EMBL; AB086357; BAC00825.1; -
CC EMBL; AK027669; BAB55281.1; -
CC EMBL; AK075033; BAC11365.1; -
CC EMBL; BC003533; AAH03533.1; -
CC EMBL; BC009383; AAH09383.1; -
CC HSSP; P00750; 1PK2.
CC Genew; HGNC:18797; KREMEN2.
CC InterPro; IPR000859; CUB domain.
CC InterPro; IPR000001; Kringle.
CC InterPro; IPR002889; WSC.
CC Pfam; PF00431; CUB; 1.
CC Pfam; PF00051; kringle; 1.
CC Pfam; PF01822; WSC; 1.
CC PRINTS; PR00018; KRINGLE.
CC ProDom; PD000395; Kringle; 1.
CC SMART; SM00042; CUB; 1.
CC SMART; SM00130; KR; 1.
CC SMART; SM00321; WSC; 1.
CC PROSITE; PS01180; CUB; 1.
CC PROSITE; PS00021; KRINGLE; 1.
CC PROSITE; PS00070; KRINGLE; 2; 1.
CC Wnt signaling pathway; Glycoprotein; Kringle; Signal; Transmembrane;
KW Alternative splicing.

Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: PLASMIN DISSOLVES THE FIBRIN OF BLOOD CLOTS AND ACTS AS A PROTEOLYTIC FACTOR IN A VARIETY OF OTHER PROCESSES INCLUDING EMBRYONIC DEVELOPMENT, TISSUE REMODELING, TUMOR INVASION AND INFLAMMATION; IN OVULATION IT WEAKENS THE WALLS OF THE GRAAFIAN FOLLICLE. IT ACTIVATES THE UROKINASE-TYPE PLASMINOGEN ACTIVATOR, COLLAGENASES AND SEVERAL COMPLEMENT ZYMOGENS, SUCH AS C1 AND C5. IT CLEAVES FIBRIN, FIBRONECTIN, THROMBOSPONDIN, LAMININ AND VON WILLEBRAND FACTOR.

-!- CATALYTIC ACTIVITY: Preferential cleavage: Lys-|-Xaa > Arg-|-Xaa; higher selectivity than trypsin. Converts fibrin into soluble products.

-!- ENZYME REGULATION: CONVERTED INTO PLASMIN BY PLASMINOGEN ACTIVATORS, BOTH PLASMINOGEN AND ITS ACTIVATOR BEING BOUND TO FIBRIN. CANNOT BE ACTIVATED WITH STREPTOKINASE.

-!- MISCELLANEOUS: PLASMIN IS INACTIVATED BY ALPHA-2-ANTIPLASMIN IMMEDIATELY AFTER DISSOCIATION FROM THE CLOT.

-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMINOGEN SUBFAMILY.

-!- SIMILARITY: Contains 5 kringle domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; U33171; AAC48717.1; -
PIR; I46260; I46260.
HSP; P00747; 1PMK.
MEROPS; S01.233; -
InterPro; IPR001314; Chymotrypsin.
InterPro; IPR000001; Kringle.
InterPro; IPR003014; PAN.
InterPro; IPR003609; Pan_app.
InterPro; IPR003966; Prothrombin.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00051; kringle; 5.
Pfam; PF00024; PAN; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00722; CHYMOTRYPSIN.
PRINTS; PR00018; KRINGLE.
PRINTS; PR01505; PROTHROMBIN.
ProDom; PD000395; Kringle; 5.
SMART; SM00130; KR; 5.
SMART; SM00473; PAN AP; 1.
SMART; SM00020; Tryp_SPC; 1.
PROSITE; PS00021; KRINGLE_1; 5.
PROSITE; PS00070; KRINGLE_2; 5.
PROSITE; PS0240; TRYPSIN_DOM; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_SER; 1.
Hydrolase; Serine protease; Plasma; Glycoprotein; Fibrinolysis; Tissue remodeling; Blood coagulation; kringle; Zymogen; Repeat; Signal.
Signal.
FT CHAIN 1 19 BY SIMILARITY.
FT CHAIN 20 810 PLASMINOGEN.
FT CHAIN 20 582 PLASMIN HEAVY CHAIN A (BY SIMILARITY).
FT CHAIN 583 810 PLASMIN LIGHT CHAIN B (BY SIMILARITY).
FT DOMAIN 583 810 SERINE PROTEASE.
FT DOMAIN 103 181 KRINGLE 1.
FT DOMAIN 185 262 KRINGLE 2.
FT DOMAIN 275 352 KRINGLE 3.
FT DOMAIN 379 456 KRINGLE 4.
FT DOMAIN 482 561 KRINGLE 5.
FT ACT_SITE 622 622 CHARGE RELAY SYSTEM.
FT ACT_SITE 665 665 CHARGE RELAY SYSTEM.
FT ACT_SITE 760 760 CHARGE RELAY SYSTEM.
FT CARBOHYD 339 339 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 810 AA; 90902 MW; 8E75780946017A16 CRC64;

Query Match 29.1%; Score 148.5; DB 1; Length 810;

Best Local Similarity 41.5%; Pred. No. 1.9e-09;
Matches 27; Conservative 7; Mismatches 28; Indels 3; Gaps 1;
Qy 3 CYEGNGHYRGKASTDTMGKPCLPWNSATVLCQTYHAHRSDALQLGLGKHNYCRNPDNR 62
Db 379 CYQNGQTYRGTSSTITGKCCQPWTSMPHRHSKTPENYPADLTM--NYCRNPDGDK 435
Qy 63 RPWCY 67
Db 436 GPWCY 440

Search completed: December 3, 2003, 14:40:04
Job time : 3.82759 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:33:02 ; Search time 13.8621 Seconds
(without alignments)
1007.637 Million cell updates/sec

Title: US-09-880-503-1

Perfect score: 510

Sequence: 1 KTCVEGNHGFYRGKASTDTM.....QVGLKPLVQRCMVHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	510	100.0	88	AAE16542	Human urokinase-ty
2	510	100.0	96	AAE16550	Human uPA kringle
3	510	100.0	135	AAE16545	Human urokinase-ty
4	510	100.0	143	AAE16549	Human uPA amino te
5	510	100.0	337	AAE16542	Human colon cancer
6	510	100.0	337	ABP41795	Human ovarian anti
7	510	100.0	365	AAE68854	Delta 1-46 urokina
8	510	100.0	378	AAW13635	Human prourokinase
9	510	100.0	386	AAE66266	Bifunctional uroki

10	510	100.0	389	11	AAW13636	Human prourokinase
11	510	100.0	390	16	AAE66245	Bifunctional uroki
12	510	100.0	390	16	AAE66247	Bifunctional uroki
13	510	100.0	392	16	AAE66254	Bifunctional uroki
14	510	100.0	392	16	AAE66255	Bifunctional uroki
15	510	100.0	392	16	AAE66256	Bifunctional uroki
16	510	100.0	392	16	AAE66257	Bifunctional uroki
17	510	100.0	392	16	AAE66258	Bifunctional uroki
18	510	100.0	392	16	AAE66259	Bifunctional uroki
19	510	100.0	392	16	AAE66260	Bifunctional uroki
20	510	100.0	392	16	AAE66261	Bifunctional uroki
21	510	100.0	392	16	AAE66263	Bifunctional uroki
22	510	100.0	393	16	AAE66264	Bifunctional uroki
23	510	100.0	393	16	AAE66244	Bifunctional uroki
24	510	100.0	393	16	AAE66249	Bifunctional uroki
25	510	100.0	393	16	AAE66250	Bifunctional uroki
26	510	100.0	393	16	AAE66251	Bifunctional uroki
27	510	100.0	393	16	AAE66252	Bifunctional uroki
28	510	100.0	393	16	AAE66253	Bifunctional uroki
29	510	100.0	393	17	AAE99885	M36: fibrinolytic
30	510	100.0	393	17	AAE99596	Chimeric protein M
31	510	100.0	393	17	AAE99597	Chimeric protein M
32	510	100.0	395	15	AAE47902	Pro-urokinase deri
33	510	100.0	395	16	AAE66262	Bifunctional uroki
34	510	100.0	395	16	AAE66265	Bifunctional uroki
35	510	100.0	396	16	AAE66246	Bifunctional uroki
36	510	100.0	397	16	AAE66248	Bifunctional uroki
37	510	100.0	401	11	AAW13637	Human prourokinase
38	510	100.0	403	23	AAE16547	Human urokinase-ty
39	510	100.0	410	13	AAE23794	Prourokinase doubl
40	510	100.0	410	13	AAE23795	Prourokinase doubl
41	510	100.0	411	6	AAE50871	Sequence encoded b
42	510	100.0	411	11	AAW13634	Human native prour
43	510	100.0	411	11	AAE05117	UK-S3 as encoded b
44	510	100.0	411	11	AAE06244	Urokinase precurs
45	510	100.0	411	11	AAE07902	Human pro-urokinas

ALIGNMENTS

RESULT 1

AAE16542
ID AAE16542 standard; Protein; 88 AA.

AC AAE16542;

DT 09-APR-2002 (first entry)

XX Human urokinase-type plasminogen activator (uPA) kringle.

XX Human: urokinase-type plasminogen activator; uPA; therapy; hypertension;
stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
clotting disorder; uterine contraction disorder; respiratory disease;
male impotence; adult respiratory distress syndrome.

XX Homo sapiens.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

DR N-PSDB; AAD27075.
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -
 XX
 PS Claim 1; Fig 1A; 117pp; English.
 XX
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle.
 XX
 XX Sequence 88 AA;
 SQ
 Query Match 100.0%; Score 510; DB 23; Length 88;
 Best Local Similarity 100.0%; Pred. No. 1.1e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
 Db 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 RESULT 2
 AAE16550
 ID AAE16550 standard; Protein; 96 AA.
 AC AAE16550;
 XX
 XX 09-APR-2002 (first entry)
 DT
 DE Human uPA kringle and connecting peptide.
 XX
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; male impotence.
 XX
 OS Homo sapiens.
 XX
 XX WO200197752-A2.
 PN
 XX 27-DEC-2001.
 PD
 XX 13-JUN-2001; 2001WO-US18976.
 PF
 XX 20-JUN-2000; 2000US-212874P.
 PR
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA
 XX Cines DB, Higazi AA;
 PI
 XX WPI; 2002-122240/16.
 DR
 XX N-PSDB; AAD27083.
 XX
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -
 XX
 PS Claim 25; Fig 1I; 117pp; English.
 XX
 CC The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, tumour cell metastasis, glaucoma,
 CC diabetic retinopathy, wound healing, clotting disorder, uterine
 CC contraction disorder, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) kringle.
 XX
 XX Sequence 96 AA;
 SQ
 Query Match 100.0%; Score 510; DB 23; Length 96;
 Best Local Similarity 100.0%; Pred. No. 1.2e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
 Db 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOQTYHAHRSALQLGLGKHNYCRNPDN 60
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 RESULT 3
 AAE16545
 ID AAE16545 standard; Protein; 135 AA.
 AC AAE16545;
 XX
 XX 09-APR-2002 (first entry)
 DT
 DE Human urokinase-type plasminogen activator amino terminal fragment (ATF).
 XX
 KW Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; amino terminal fragment; ATF;
 KW male impotence.
 XX
 OS Homo sapiens.
 XX
 XX WO200197752-A2.
 PN
 XX 27-DEC-2001.
 PD
 XX 13-JUN-2001; 2001WO-US18976.
 PF
 XX 20-JUN-2000; 2000US-212874P.
 PR
 XX (UYPE-) UNIV PENNSYLVANIA.
 PA
 XX Cines DB, Higazi AA;
 PI
 XX WPI; 2002-122240/16.
 DR
 XX N-PSDB; AAD27078.
 XX
 XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,

PT comprising domains from urokinase-type plasminogen activator -
 XX
 PS Claim 11; Fig 1D; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, wound healing, clotting disorder, uterine
 CC diabetic retinopathy, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) amino
 CC terminal fragment (ATF).

XX SQ Sequence 135 AA;

Query Match 100.0%; Score 510; DB 23; Length 135;
 Best Local Similarity 100.0%; Pred. NO. 1.7e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 DB 48 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 4

AAE16549
 ID AAE16549 standard; Protein; 143 AA.

AC AAE16549;

DT 09-APR-2002 (first entry)

XX Human uPA amino terminal fragment (ATF) and connecting peptide.

DE Human; urokinase-type plasminogen activator; uPA; therapy; hypertension;
 KW stroke; hypotension; atherosclerosis; heart attack; thrombotic disorder;
 KW microvascular occlusion; angiogenic disorder; pulmonary fibrosis; asthma;
 KW tumour cell metastasis; glaucoma; diabetic retinopathy; wound healing;
 KW clotting disorder; uterine contraction disorder; respiratory disease;
 KW adult respiratory distress syndrome; amino terminal fragment; ATF;
 KW male impotence.

XX Homo sapiens.

XX WO200197752-A2.

XX 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US18976.

XX 20-JUN-2000; 2000US-212874P.

XX (UYPE-) UNIV PENNSYLVANIA.

XX Cines DB, Higazi AA;

XX WPI; 2002-122240/16.

XX N-PSDB; AAD27082.

XX Composition for modulating muscle cell and tissue contractility for
 PT treating atherosclerosis, asthma, hypertension, glaucoma, impotence,
 PT comprising domains from urokinase-type plasminogen activator -

XX Claim 24; Fig 1H; 117pp; English.

XX The invention relates to a composition comprising one or more domains of
 CC urokinase-type plasminogen activator (uPA). The composition is used to
 CC modulate the contractility and angiogenic activity of a mammalian muscle,
 CC endothelial cell or tissue. The composition is used for treating stroke,
 CC hypotension, hypertension, atherosclerosis, heart attack, microvascular
 CC occlusions, thrombotic microangiopathies, surgically induced thrombotic
 CC disorders, angiogenic disorders, pulmonary fibrosis, asthma, tumour cell
 CC invasion, tumour cell angiogenesis, wound healing, clotting disorder, uterine
 CC diabetic retinopathy, male impotence, respiratory disease or condition
 CC such as asthma, adult respiratory distress syndrome, primary pulmonary
 CC hypertension, microvascular thrombotic occlusion, and a disorder
 CC associated with chronic intrapulmonary fibrin formation. The present
 CC sequence is human urokinase-type plasminogen activator (uPA) amino
 CC terminal fragment (ATF) and connecting peptide.

XX SQ Sequence 143 AA;

Query Match 100.0%; Score 510; DB 23; Length 143;
 Best Local Similarity 100.0%; Pred. NO. 1.8e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 60
 DB 48 KTCYEGNGHFYRGKASTDTWGRPCLPWSATVLQOQTYHAHRSDALQLGLGKHNYCRNPDN 107
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 DB 108 RRRPWCYVQVGLKPLVQECMVHDCADGK 135

RESULT 5

AAAG75492
 ID AAG75492 standard; Protein; 337 AA.

AC AAG75492;

DT 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:6256.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; chromosome 10.

XX Homo sapiens.

XX WO200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000WO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Birse CE, Rosen CA;

XX WPI; 2001-235357/24.

XX N-PSDB; AAH34897.

XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 7707-7708; 9803pp; English.

XX AAH32943 to AAH37195 and AAG773514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon

CC cancer antigens have cytostatic activity and can be used in gene
CC therapy and vaccine production. N and P may be used in the prevention,
CC diagnosis and treatment of diseases associated with inappropriate p
CC expression. For example, N and P may be used to treat disorders
CC associated with decreased expression by rectifying mutations or deletions
CC in a patient's genome that affect the activity of p by expressing
CC inactive proteins or to supplement the patients own production of p.
CC Additionally, N may be used to produce the colon cancer-associated Ps,
CC by inserting the nucleic acids into a host cell and culturing the cell
CC to express the proteins. N and P can be used in the prevention, diagnosis
CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC and AAH77789 represent sequences used in the exemplification of the
CC present invention.
CC N.B. Pages 665 to 682 and page 7053 of the sequence listing were
CC missing at time of publication, meaning no sequences are present for
CC SEQ ID NO:1027 to 1052, 7921 and 7922.

XX SQ Sequence 337 AA;

Query Match 100.0%; Score 510; DB 22; Length 337;
Best Local Similarity 100.0%; Pred. No. 4e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60

Db 74 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 133

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

Db 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 6

ABP41795

ID ABP41795 standard; Protein; 337 AA.

XX AC ABP41795;

XX DT 22-AUG-2002 (first entry)

XX DE Human ovarian antigen HVVCB79, SEQ ID NO:2927.

XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
XX KW inflammatory condition; immune disorder; blood disorder;
XX KW cardiovascular disorder; respiratory disorder; neurological disorder;
XX KW gastrointestinal disorder; urinary system disorder; drug screening;
XX KW gene therapy; chromosome mapping; forensic analysis;
XX KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX KW anti-inflammatory; gynaecological; reproductive; chromosome 10q24.

XX OS Homo sapiens.

XX FN WO200200677-A1.

XX PD 03-JAN-2002.

XX PF 07-JUN-2001; 2001WO-US18569.

XX FR 07-JUN-2000; 2000US-209467P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Birse CE, Rosen CA;

XX DR WPI; 2002-147878/19.

XX DR N-PSDB; ABQ54872.

XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g.
XX ovarian cancer), immune disorders, cardiovascular disorders and

PT neurological diseases -

XX PS Claim 11; SEQ ID NO 2927; 2922pp; English.

XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 337 AA;

Query Match 100.0%; Score 510; DB 23; Length 337;
Best Local Similarity 100.0%; Pred. No. 4e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60

Db 74 KTCYEGNGHFFYRGKASTDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 133

Qy 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88

Db 134 RRRPWCYVQVGLKPLVQECMVHDCADGK 161

RESULT 7

AAR68854

ID AAR68854 standard; protein; 365 AA.

XX AC AAR68854;

XX DT 25-MAR-2003 (updated)

XX DT 22-NOV-1995 (first entry)

XX DE Delta 1-46 urokinase.

XX KW Human; des-epidermal growth factor homologous plasminogen activator;
XX uPA; liver membrane; reduced affinity; BGF homologous; thrombosis;
XX KW thrombolytic; increased half-life; urokinase.

XX OS Homo sapiens (engineered).

XX FH Key Location/Qualifiers

XX FT Misc-difference 1

XX FT /note= "amino acids 1-46 of wild-type urokinase
XX have been deleted"

XX PN US5376547-A.

PD XX 27-DEC-1994.
 PF XX 29-JAN-1988; 88US-0150267.
 PR XX 30-JAN-1987; 87US-0008795.
 PR XX 29-JAN-1988; 88US-0150267.
 XX XX (AMHP) AMERICAN HOME PROD CORP.
 XX XX Hung PP, Kalyan NK, Lee SL;
 XX XX WPI; 1995-043464/06.
 DR XX New modified plasminogen activator cpds. - having regions removed
 PT XX to reduce affinity for liver membranes and increase circulation
 PT XX half-life.
 XX XX Claim 1; Page ?; 26pp; English.
 PS XX Amino acid residues 1-46 contain the EGF region of human urokinase.
 CC XX Deletion of this region results in a plasminogen activator with
 CC XX reduced affinity for liver cell membranes; the mutant protein is
 CC XX not cleared from the circulation as rapidly as is wild-type tPA.
 CC XX The specification only gives the sequence around the deletion and
 CC XX not the full-length sequence of "delta 1-46 urokinase"; the
 CC XX sequence in AAR68854 has been obtained by amending a previously
 CC XX disclosed wild-type human urokinase sequence (from W09501427)
 CC XX according to the description given in Example 3.
 CC XX (Updated on 25-MAR-2003 to correct PF field.)
 XX XX
 SQ Sequence 365 AA;
 Query Match 100.0%; Score 510; DB 16; Length 365;
 Best Local Similarity 100.0%; Pred. No. 4.3e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60
 DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 61
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
 RESULT 8
 AAW13635
 ID AAW13635 standard; Protein; 378 AA.
 XX AC AAW13635;
 XX DT 04-JUN-1997 (first entry)
 XX DE Human prourokinase variant lacking entire EGF domain.
 XX KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
 KW KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
 XX OS Homo sapiens.
 XX OS Synthetic.
 FH Key Location/Qualifiers
 FT Region 1..9
 FT /note= "residues 1-9 of native hPUK"
 FT Region 10..378
 FT /note= "residues 43-411 of native hPUK"
 FT Misc-difference 144
 FT /note= "corresponds to TAC codon"
 XX EP398361-A.
 PN 22-NOV-1990.
 PD XX

PF 18-MAY-1990; 90EP-0109472.
 XX 22-FEB-1990; 90JP-0042020.
 PR 18-MAY-1989; 89JP-0126433.
 PR 03-JUL-1986; 86JP-0156936.
 PR 18-FEB-1987; 87JP-0036495.
 PR 18-MAY-1989; 89JP-0126434.
 XX (GREC) GREEN CROSS CORP.
 XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
 PI Morita M, Tanabe T;
 XX WPI; 1990-350146/47.
 DR N-PSDB; AAT61672.
 XX Human pro-urokinase variants - deficient in loop regions of
 PT epidermal growth factor, showing long blood half-life, as
 PT fibrinolytic agent
 XX Claim 1; Page -; 22pp; English.
 PS New variants of human prourokinase (hPUK) comprise a hPUK deficient
 CC in (i) at least part of the first loop region of the epidermal growth
 CC factor (EGF) domain; (ii) at least part of the first loop and at
 CC least part of the second loop; or (iii) at least part of the third
 CC loop. The hPUK variants show an increased blood half-life comparable
 CC to that of the whole EGF domain-deficient hPUK variant and urokinase
 CC while retaining the same properties as those of hPUK. They have
 CC potent thrombolytic activity and very little tendency to cause
 CC spontaneous bleeding. The present sequence represents a specific variant
 CC of hPUK which lacks the entire EGF domain; the sequence does not
 CC appear in the specification and has been created using the
 CC wild-type hPUK sequence and the junction sequence after
 CC deletion, both of which are given (in Fig 1 and in Fig 2(3),
 CC respectively).
 XX Sequence 378 AA;
 Query Match 100.0%; Score 510; DB 11; Length 378;
 Best Local Similarity 100.0%; Pred. No. 4.5e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 60
 DB 15 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRS DALQLGLGKHNYCRNPDN 74
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 DB 75 RRRPWCYVQVGLKPLVQECMVHDCADGK 102
 RESULT 9
 AAR66266
 ID AAR66266 standard; protein; 386 AA.
 XX AC AAR66266;
 XX DT 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX DE Bifunctional urokinase variant M33.
 XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 KW KW urokinase; variant; mutein.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Region 1..365
 FT /label= M4
 FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85

FT Disulfide-bond 25...67
 FT Disulfide-bond 56...80
 FT Disulfide-bond 102...1233
 FT Disulfide-bond 143...159
 FT Disulfide-bond 151...222
 FT Disulfide-bond 247...316
 FT Disulfide-bond 279...295
 FT Disulfide-bond 306...334
 FT Region 365...366
 FT /label= X1
 FT /note= "peptide bond"
 FT /label= Y1
 PN DE4323754-Cl.
 XX
 PD 01-DEC-1994.
 XX
 XX 15-JUL-1993; 93DE-4323754.
 XX
 PR 15-JUL-1993; 93DE-4323754.
 XX
 XX (CHEF) GRUENTHAL GMBH.
 PA
 XX Heinzl-Wieland R, Saunders DU, Schneider J, Steffens GJ;
 PI Wndndt S;
 XX WPI; 1995-015191/03.
 DR
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX
 PS Example 1; Page 11 and Fig 1; 34pp; German.
 XX
 CC Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 386 AA;
 Query Match 100.0%; Score 510; DB 16; Length 386;
 Best Local Similarity 100.0%; Pred. No. 4.6e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
 Db
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
 RESULT 10
 ID AAW13636
 XX AAW13636 standard; Protein; 389 AA.
 AC AAW13636;
 XX
 DT 04-JUN-1997 (first entry)
 XX
 DE Human prourokinase variant lacking EGF domain loops 1 and 2.
 KW Human; prourokinase; hPUK; variant; half-life; increase; EGF;
 KW epidermal growth factor domain; deletion; thrombolysis; fibrinolysis.
 XX

OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key
 FT Region Location/Qualifiers
 FT 1..10
 FT /note= "residues 1-10 of native hPUK"
 FT Region 11..389
 FT /note= "residues 33-411 of native hPUK"
 FT Misc-difference 155
 FT /note= "corresponds to TAC codon"
 XX
 XX EP398361-A.
 XX
 PD 22-NOV-1990.
 XX
 PF 18-MAY-1990; 90EP-0109472.
 XX
 PR 22-FEB-1990; 90JP-0042020.
 PR 18-MAY-1989; 89JP-0126433.
 PR 03-JUL-1986; 86JP-0156936.
 PR 18-FEB-1987; 87JP-0036495.
 PR 18-MAY-1989; 89JP-0126434.
 XX
 XX (GRC) GREEN CROSS CORP.
 PA
 XX Airmura H, Amatsuji Y, Hirose M, Kasai S, Kawabe H;
 PI Morita M, Tanabe T;
 XX
 DR WPI; 1990-350146/47.
 DR N-PSDB; AAT61673.
 XX
 XX Human pro-urokinase variants - deficient in loop regions of
 PT epidermal growth factor, showing long blood half-life, as
 PT fibrinolytic agent
 XX
 PS Claim 6; Page -; 22pp; English.
 XX
 CC New variants of human prourokinase (hPUK) comprise a hPUK deficient
 CC in (i) at least part of the first loop region of the epidermal growth
 CC factor (EGF) domain; (ii) at least part of the first loop and at
 CC least part of the second loop; or (iii) at least part of the third
 CC loop. The hPUK variants show an increased blood half-life comparable
 CC to that of the whole EGF domain-deficient hPUK variant and urokinase
 CC while retaining the same properties as those of hPUK. They have
 CC potent thrombolytic activity and very little tendency to cause
 CC spontaneous bleeding. The present sequence represents a specific
 CC variant of hPUK which lacks loops 1 and 2 of the EGF domain; the
 CC sequence does not appear in the specification and has been created
 CC using the wild-type hPUK sequence and the junction sequence after
 CC deletion, both of which are given (in Fig 1 and on page 8,
 CC respectively).
 SQ Sequence 389 AA;
 Query Match 100.0%; Score 510; DB 11; Length 389;
 Best Local Similarity 100.0%; Pred. No. 4.6e-41;
 Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 60
 Db 26 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLQOTYHAHRSDALQLGLGKHNYCRNPDN 85
 QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
 Db 86 RRRPWCYVQVGLKPLVQECMVHDCADGK 113
 RESULT 11
 ID AAR66245
 XX AAR66245 standard; protein; 390 AA.
 AC AAR66245;
 XX

```
DT 25-MAR-2003 (updated)
DE 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M12.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT /label= X1
FT /label= Y1
XX
PN DB4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heintel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 390 AA;
Query Match 100.0%; Score 510; DB 16; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPN 61
QY 61 RRRPWCYVQGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQGLKPLVQECMVHDCADGK 89
```

```
RESULT 12
AAR66247
ID AAR66247 standard; protein; 390 AA.
XX
XX AAR66247;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M14.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT /label= X1
FT /label= Y1
XX
PN DE4323754-C1.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heintel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 390 AA;
Query Match 100.0%; Score 510; DB 16; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNCRPN 60
```

4 KTCYEGNGHFGKASDTDTWGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 13
AAR66254
ID AAR66254 standard; protein; 392 AA.
AC AAR66254;
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M21.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.

Key Location/Qualifiers
Region 1..365
/label= M4
/note= "unglycosylated prourokinase (Ser47-Leu411)"
Disulfide-bond 4..82
Disulfide-bond 25..87
Disulfide-bond 56..80
Disulfide-bond 102..233
Disulfide-bond 143..159
Disulfide-bond 151..222
Disulfide-bond 247..316
Disulfide-bond 279..295
Disulfide-bond 306..334
Region 366..371
/label= X1
Region 372..392
/label= Y1

DE4323754-C1.
O1-DEC-1994.
15-JUL-1993; 93DE-4323754.
15-JUL-1993; 93DE-4323754.
(CHEF) GRUENENTHAL GMBH.
Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
Wendt S;
WPI; 1995-015191/03.

New bifunctional urokinase derivs and related plasmids - with improved fibrinolytic and thrombin inhibiting activities, for treating cardiac and cerebral infarct, pulmonary embolism, etc
Example 1; Page 10 and Fig 1; 34pp; German.
Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266 are specific examples of such derivs which have both improved fibrinolytic and thrombin-inhibiting activities compared to known plasminogen activators or thrombin inhibitors. The proteins are useful as thrombolytic agents, e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac and cerebral infarction and pulmonary embolism.
Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 1 KTCYEGNGHFGKASDTDTWGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNCRNPDN 60
Db 2 KTCYEGNGHFGKASDTDTWGRPCLPWN SATVLQOYTHAHRSDALQGLGKHNCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 14
AAR66255
ID AAR66255 standard; protein; 392 AA.
AC AAR66255;
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX Bifunctional urokinase variant M22.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX Synthetic.

Key Location/Qualifiers
Region 1..365
/label= M4
/note= "unglycosylated prourokinase (Ser47-Leu411)"
Disulfide-bond 4..82
Disulfide-bond 25..87
Disulfide-bond 56..80
Disulfide-bond 102..233
Disulfide-bond 143..159
Disulfide-bond 151..222
Disulfide-bond 247..316
Disulfide-bond 279..295
Disulfide-bond 306..334
Region 366..371
/label= X1
Region 372..392
/label= Y1

DE4323754-C1.
O1-DEC-1994.
15-JUL-1993; 93DE-4323754.
15-JUL-1993; 93DE-4323754.
(CHEF) GRUENENTHAL GMBH.
Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
Wendt S;
WPI; 1995-015191/03.

New bifunctional urokinase derivs and related plasmids - with improved fibrinolytic and thrombin inhibiting activities, for treating cardiac and cerebral infarct, pulmonary embolism, etc
Example 1; Page 10 and Fig 1; 34pp; German.
Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1 are claimed (see features table). Sequences AAR6244-R66266 are specific examples of such derivs. which have both improved

```
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. NO. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 15
AAR66256
ID AAR66256 standard; protein; 392 AA.
XX
AC AAR66256;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M23.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FT Key Location/Qualifiers
FT Region 1..365 /label= M4
FT Disulfide-bond 4..85 /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..392 /label= Y1
XX
DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Wnendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
```

```
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. NO. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLTQTYHAHRSALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 16
AAR66257
ID AAR66257 standard; protein; 392 AA.
XX
AC AAR66257;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M24.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FT Key Location/Qualifiers
FT Region 1..365 /label= M4
FT Disulfide-bond 4..85 /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..371
FT Region /label= X1
FT Region 372..392 /label= Y1
XX
DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
```

UR WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX (CHEP) GRUENENTHAL GMBH.
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PS Wnendt S;
 XX WPI; 1995-015191/03.
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 392 AA;
 XX Query Match 100.0%; Score 510; DB 16; Length 392;
 XX Best Local Similarity 100.0%; Pred. No. 4.6e-41;
 XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTVLTQTYHAHRSALQLGLGKHYCRNPDN 60
 DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTVLTQTYHAHRSALQLGLGKHYCRNPDN 61
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
 DB 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89
 RESULT 17
 ID AAR66258
 XX AAR66258 standard; protein: 392 AA.
 XX AC AAR66258;
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M25.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 XX urokinase; variant; mutein.
 XX Synthetic.
 XX Key Location/Qualifiers
 XX Region 1..365
 XX /label= M4
 XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 XX /label= X1
 XX /label= Y1
 XX 366..371
 XX 372..392
 XX 393..411
 XX 412..434
 XX 435..457
 XX 458..480
 XX 481..503
 XX 504..526
 XX 527..549
 XX 550..572
 XX 573..595
 XX 596..618
 XX 619..641
 XX 642..664
 XX 665..687
 XX 688..710
 XX 711..733
 XX 734..756
 XX 757..779
 XX 780..802
 XX 803..825
 XX 826..848
 XX 849..871
 XX 872..894
 XX 895..917
 XX 918..940
 XX 941..963
 XX 964..986
 XX 987..1009
 XX 1010..1032
 XX 1033..1055
 XX 1056..1078
 XX 1079..1101
 XX 1102..1124
 XX 1125..1147
 XX 1148..1170
 XX 1171..1193
 XX 1194..1216
 XX 1217..1239
 XX 1240..1262
 XX 1263..1285
 XX 1286..1308
 XX 1309..1331
 XX 1332..1354
 XX 1355..1377
 XX 1378..1400
 XX 1401..1423
 XX 1424..1446
 XX 1447..1469
 XX 1470..1492
 XX 1493..1515
 XX 1516..1538
 XX 1539..1561
 XX 1562..1584
 XX 1585..1607
 XX 1608..1630
 XX 1631..1653
 XX 1654..1676
 XX 1677..1699
 XX 1700..1722
 XX 1723..1745
 XX 1746..1768
 XX 1769..1791
 XX 1792..1814
 XX 1815..1837
 XX 1838..1860
 XX 1861..1883
 XX 1884..1906
 XX 1907..1929
 XX 1930..1952
 XX 1953..1975
 XX 1976..1998
 XX 1999..2021
 XX 2022..2044
 XX 2045..2067
 XX 2068..2090
 XX 2091..2113
 XX 2114..2136
 XX 2137..2159
 XX 2160..2182
 XX 2183..2205
 XX 2206..2228
 XX 2229..2251
 XX 2252..2274
 XX 2275..2297
 XX 2298..2320
 XX 2321..2343
 XX 2344..2366
 XX 2367..2389
 XX 2390..2412
 XX 2413..2435
 XX 2436..2458
 XX 2459..2481
 XX 2482..2504
 XX 2505..2527
 XX 2528..2550
 XX 2551..2573
 XX 2574..2596
 XX 2597..2619
 XX 2620..2642
 XX 2643..2665
 XX 2666..2688
 XX 2689..2711
 XX 2712..2734
 XX 2735..2757
 XX 2758..2780
 XX 2781..2803
 XX 2804..2826
 XX 2827..2849
 XX 2850..2872
 XX 2873..2895
 XX 2896..2918
 XX 2919..2941
 XX 2942..2964
 XX 2965..2987
 XX 2988..3010
 XX 3011..3033
 XX 3034..3056
 XX 3057..3079
 XX 3080..3102
 XX 3103..3125
 XX 3126..3148
 XX 3149..3171
 XX 3172..3194
 XX 3195..3217
 XX 3218..3240
 XX 3241..3263
 XX 3264..3286
 XX 3287..3309
 XX 3310..3332
 XX 3333..3355
 XX 3356..3378
 XX 3379..3401
 XX 3402..3424
 XX 3425..3447
 XX 3448..3470
 XX 3471..3493
 XX 3494..3516
 XX 3517..3539
 XX 3540..3562
 XX 3563..3585
 XX 3586..3608
 XX 3609..3631
 XX 3632..3654
 XX 3655..3677
 XX 3678..3700
 XX 3701..3723
 XX 3724..3746
 XX 3747..3769
 XX 3770..3792
 XX 3793..3815
 XX 3816..3838
 XX 3839..3861
 XX 3862..3884
 XX 3885..3907
 XX 3908..3930
 XX 3931..3953
 XX 3954..3976
 XX 3977..3999
 XX 4000..4022
 XX 4023..4045
 XX 4046..4068
 XX 4069..4091
 XX 4092..4114
 XX 4115..4137
 XX 4138..4160
 XX 4161..4183
 XX 4184..4206
 XX 4207..4229
 XX 4230..4252
 XX 4253..4275
 XX 4276..4298
 XX 4299..4321
 XX 4322..4344
 XX 4345..4367
 XX 4368..4390
 XX 4391..4413
 XX 4414..4436
 XX 4437..4459
 XX 4460..4482
 XX 4483..4505
 XX 4506..4528
 XX 4529..4551
 XX 4552..4574
 XX 4575..4597
 XX 4598..4620
 XX 4621..4643
 XX 4644..4666
 XX 4667..4689
 XX 4690..4712
 XX 4713..4735
 XX 4736..4758
 XX 4759..4781
 XX 4782..4804
 XX 4805..4827
 XX 4828..4850
 XX 4851..4873
 XX 4874..4896
 XX 4897..4919
 XX 4920..4942
 XX 4943..4965
 XX 4966..4988
 XX 4989..5011
 XX 5012..5034
 XX 5035..5057
 XX 5058..5080
 XX 5081..5103
 XX 5104..5126
 XX 5127..5149
 XX 5150..5172
 XX 5173..5195
 XX 5196..5218
 XX 5219..5241
 XX 5242..5264
 XX 5265..5287
 XX 5288..5310
 XX 5311..5333
 XX 5334..5356
 XX 5357..5379
 XX 5380..5402
 XX 5403..5425
 XX 5426..5448
 XX 5449..5471
 XX 5472..5494
 XX 5495..5517
 XX 5518..5540
 XX 5541..5563
 XX 5564..5586
 XX 5587..5609
 XX 5610..5632
 XX 5633..5655
 XX 5656..5678
 XX 5679..5701
 XX 5702..5724
 XX 5725..5747
 XX 5748..5770
 XX 5771..5793
 XX 5794..5816
 XX 5817..5839
 XX 5840..5862
 XX 5863..5885
 XX 5886..5908
 XX 5909..5931
 XX 5932..5954
 XX 5955..5977
 XX 5978..6000
 XX 6001..6023
 XX 6024..6046
 XX 6047..6069
 XX 6070..6092
 XX 6093..6115
 XX 6116..6138
 XX 6139..6161
 XX 6162..6184
 XX 6185..6207
 XX 6208..6230
 XX 6231..6253
 XX 6254..6276
 XX 6277..6299
 XX 6300..6322
 XX 6323..6345
 XX 6346..6368
 XX 6369..6391
 XX 6392..6414
 XX 6415..6437
 XX 6438..6460
 XX 6461..6483
 XX 6484..6506
 XX 6507..6529
 XX 6530..6552
 XX 6553..6575
 XX 6576..6598
 XX 6599..6621
 XX 6622..6644
 XX 6645..6667
 XX 6668..6690
 XX 6691..6713
 XX 6714..6736
 XX 6737..6759
 XX 6760..6782
 XX 6783..6805
 XX 6806..6828
 XX 6829..6851
 XX 6852..6874
 XX 6875..6897
 XX 6898..6920
 XX 6921..6943
 XX 6944..6966
 XX 6967..6989
 XX 6990..7012
 XX 7013..7035
 XX 7036..7058
 XX 7059..7081
 XX 7082..7104
 XX 7105..7127
 XX 7128..7150
 XX 7151..7173
 XX 7174..7196
 XX 7197..7219
 XX 7220..7242
 XX 7243..7265
 XX 7266..7288
 XX 7289..7311
 XX 7312..7334
 XX 7335..7357
 XX 7358..7380
 XX 7381..7403
 XX 7404..7426
 XX 7427..7449
 XX 7450..7472
 XX 7473..7495
 XX 7496..7518
 XX 7519..7541
 XX 7542..7564
 XX 7565..7587
 XX 7588..7610
 XX 7611..7633
 XX 7634..7656
 XX 7657..7679
 XX 7680..7702
 XX 7703..7725
 XX 7726..7748
 XX 7749..7771
 XX 7772..7794
 XX 7795..7817
 XX 7818..7840
 XX 7841..7863
 XX 7864..7886
 XX 7887..7909
 XX 7910..7932
 XX 7933..7955
 XX 7956..7978
 XX 7979..8001
 XX 8002..8024
 XX 8025..8047
 XX 8048..8070
 XX 8071..8093
 XX 8094..8116
 XX 8117..8139
 XX 8140..8162
 XX 8163..8185
 XX 8186..8208
 XX 8209..8231
 XX 8232..8254
 XX 8255..8277
 XX 8278..8300
 XX 8301..8323
 XX 8324..8346
 XX 8347..8369
 XX 8370..8392
 XX 8393..8415
 XX 8416..8438
 XX 8439..8461
 XX 8462..8484
 XX 8485..8507
 XX 8508..8530
 XX 8531..8553
 XX 8554..8576
 XX 8577..8599
 XX 8600..8622
 XX 8623..8645
 XX 8646..8668
 XX 8669..8691
 XX 8692..8714
 XX 8715..8737
 XX 8738..8760
 XX 8761..8783
 XX 8784..8806
 XX 8807..8829
 XX 8830..8852
 XX 8853..8875
 XX 8876..8898
 XX 8899..8921
 XX 8922..8944
 XX 8945..8967
 XX 8968..8990
 XX 8991..9013
 XX 9014..9036
 XX 9037..9059
 XX 9060..9082
 XX 9083..9105
 XX 9106..9128
 XX 9129..9151
 XX 9152..9174
 XX 9175..9197
 XX 9198..9220
 XX 9221..9243
 XX 9244..9266
 XX 9267..9289
 XX 9290..9312
 XX 9313..9335
 XX 9336..9358
 XX 9359..9381
 XX 9382..9404
 XX 9405..9427
 XX 9428..9450
 XX 9451..9473
 XX 9474..9496
 XX 9497..9519
 XX 9520..9542
 XX 9543..9565
 XX 9566..9588
 XX 9589..9611
 XX 9612..9634
 XX 9635..9657
 XX 9658..9680
 XX 9681..9703
 XX 9704..9726
 XX 9727..9749
 XX 9750..9772
 XX 9773..9795
 XX 9796..9818
 XX 9819..9841
 XX 9842..9864
 XX 9865..9887
 XX 9888..9910
 XX 9911..9933
 XX 9934..9956
 XX 9957..9979
 XX 9980..10000

XX 15-JUL-1993; 93DE-4323754.
 XX 15-JUL-1993; 93DE-4323754.
 XX (CHEP) GRUENENTHAL GMBH.
 XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
 PI Wnendt S;
 XX WPI; 1995-015191/03.
 XX New bifunctional urokinase derivs and related plasmids - with
 PT improved fibrinolytic and thrombin inhibiting activities, for
 PT treating cardiac and cerebral infarct, pulmonary embolism, etc
 XX Example 1; Page 11 and Fig 1; 34pp; German.
 XX Bifunctional urokinase derivatives corresponding to the formula
 CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
 CC are specific examples of such derivs. which have both improved
 CC fibrinolytic and thrombin-inhibiting activities, compared to known
 CC plasminogen activators or thrombin inhibitors. The proteins are
 CC useful as thrombolytic agents, e.g. for treatment of arterial
 CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
 CC and pulmonary embolism.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 392 AA;
 XX Query Match 100.0%; Score 510; DB 16; Length 392;
 XX Best Local Similarity 100.0%; Pred. No. 4.6e-41;
 XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTVLTQTYHAHRSALQLGLGKHYCRNPDN 60
 DB 2 KTCYEGNGHFYRGKASTDTMGRCPLPWNSTVLTQTYHAHRSALQLGLGKHYCRNPDN 61
 QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
 DB 62 RRRPWCYVQVGLKPLVQECWVHDCADGK 89
 RESULT 18
 ID AAR66259
 XX AAR66259 standard; protein: 392 AA.
 XX AC AAR66259;
 XX 25-MAR-2003 (updated)
 DT 17-AUG-1995 (first entry)
 XX Bifunctional urokinase variant M26.
 XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
 XX urokinase; variant; mutein.
 XX Synthetic.
 XX Key Location/Qualifiers
 XX Region 1..365
 XX /label= M4
 XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
 FT Disulfide-bond 4..85
 FT Disulfide-bond 25..67
 FT Disulfide-bond 56..80
 FT Disulfide-bond 102..233
 FT Disulfide-bond 143..159
 FT Disulfide-bond 151..222
 FT Disulfide-bond 247..316
 FT Disulfide-bond 279..295
 FT Disulfide-bond 306..334
 FT Region 366..371
 XX /label= X1
 XX /label= Y1
 XX 366..371
 XX 372..392
 XX 393..411
 XX 412..434
 XX 435..457
 XX 458..480
 XX 481..503
 XX 504..526
 XX 527..549
 XX 550..572
 XX 573..595
 XX 596..618
 XX 619..641
 XX 642..664
 XX 665..687
 XX 688..710
 XX 711..733
 XX 734..756
 XX 757..779
 XX 780..802
 XX 803..825
 XX 826..848
 XX 849..871
 XX 872..894
 XX 895..917
 XX 918..940
 XX 941..963
 XX 964..986
 XX 987..1009
 XX 1010..1032
 XX 1033..1055
 XX 1056..1078
 XX 1079..1101
 XX 1102..1124
 XX 1125..1147
 XX 1148..1170
 XX 1171..1193
 XX 1194..1216
 XX 1217..1239
 XX 1240..1262
 XX 1263..1285
 XX 1286..1308
 XX 1309..1331
 XX 1332..1354
 XX 1355..1377
 XX 1378..1400
 XX 1401..1423
 XX 1424..1446
 XX 1447..1469
 XX 1470..1492
 XX 1493..1515
 XX 1516..1538
 XX 1539..1561
 XX 1562..1584
 XX 1585..1607
 XX 1608..1630
 XX 1631..1653
 XX 1654..1676
 XX 1677..1699
 XX 1700..1722
 XX 1723..1745
 XX 1746..1768
 XX 1769..1791
 XX 1792..1814
 XX 1815..1837
 XX 1838..1860
 XX 1861..1883
 XX 1884..1906
 XX 1907..1929
 XX 1930..1952
 XX 1953..1975
 XX 1976..1998
 XX 1999..2021
 XX 2022..2044
 XX 2045..2067
 XX 2068..2090
 XX 2091..2113
 XX 2114..2136
 XX 2137..2159
 XX 2160..2182
 XX 2183..2205
 XX 2206..2228
 XX 2229..2251
 XX 2252..2274
 XX 2275..2297
 XX 2298..2320
 XX 2321..2343
 XX 2344..2366
 XX 2367..2389
 XX 2390..2412
 XX 2413..2435
 XX 2436..2458
 XX 2459..2481
 XX 2482..2504
 XX 2505..2527
 XX 2528..2550
 XX 2551..2573
 XX 2574..2596
 XX 2597..2619
 XX 2620..2642
 XX 2643..2665
 XX 2666..2688
 XX 2689..2711
 XX 2712..2734
 XX 2735..2757
 XX 2758..2780
 XX 2781..2803
 XX 2804..2826
 XX 2827..2849
 XX 2850..2872
 XX 2873..2895
 XX 2896..2918
 XX 2919..2941
 XX 2942..2964
 XX 2965..2987
 XX 2988..3010
 XX 3011..3033
 XX 3034..3056
 XX 3057..3079
 XX 3080..3102
 XX 3103..3125
 XX 3126..3148
 XX 3149..3171
 XX 3172..3194
 XX 3195..3217
 XX 3218..3240
 XX 3241..3263
 XX 3264..3286
 XX 3287..3309
 XX 3310..3332
 XX 3333..3355
 XX 3356..3378
 XX 3379..3401
 XX 3402..3424
 XX 3425..3447
 XX 3448..3470
 XX 3471..3493
 XX 3494..3516
 XX 3517..3539
 XX 3540..3562
 XX 3563..3585
 XX 3586..3608
 XX 3609..3631
 XX 3632..3654
 XX 3655..3677
 XX 3678..3700
 XX 3701..3723
 XX 3724..3746
 XX 3747..3769
 XX 3770..3792
 XX 3793..3815
 XX 3816..3838
 XX 3839..3861
 XX 3862..3884
 XX 3885..3907
 XX 3908..3930
 XX 3931..3953
 XX 3954..3976
 XX 3977..3999
 XX 4000..4022
 XX 4023..4045
 XX 4046..4068
 XX 4069..4091
 XX 4092..4114
 XX 4115..4137
 XX 4138..4160
 XX 4161..4183
 XX 4184..4206
 XX 4207..4229
 XX 4230..4252
 XX 4253..4275
 XX 4276..4298
 XX 4299..4321
 XX 4322..4344
 XX 4345..4367
 XX 4368..4390
 XX 4391..4413
 XX 4414..4436
 XX 4437..4459
 XX 4460..4482
 XX 4483..4505
 XX 4506..4528
 XX 4529..4551
 XX 4552..4574
 XX 4575..4597
 XX 4598..4620
 XX 4621..4643
 XX 4644..4666
 XX 4667..4689
 XX 4690..4712
 XX 4713..4735
 XX 4736..4758
 XX 4759..4781
 XX 4782..4804
 XX 4805..4827
 XX 4828..4850
 XX 4851..4873
 XX 4874..4896
 XX 4897..4919
 XX 4920..4942
 XX 4943..4965
 XX 4966..4988
 XX 4989..5011
 XX 5012..5034
 XX 5035..5057
 XX 5058..5080
 XX 5081..5103
 XX 5104..5126
 XX 5127..5149
 XX 5150..5172
 XX 5173..5195
 XX 5196..5218
 XX 5219..5241
 XX 5242..5264
 XX 5265..5287
 XX 5288..5310
 XX 5311..5333
 XX 5334..5356
 XX 5357..5379
 XX 5380..5402
 XX 5403..5425
 XX 5426..5448
 XX 5449..5471
 XX 5472..5494
 XX 5495..5517
 XX 5518..5540
 XX 5541..5563
 XX 5564..5586
 XX 5587..5609
 XX 5610..5632
 XX 5633..5655
 XX 5656..5678
 XX 5679..5701
 XX 5702..5724
 XX 5725..5747
 XX 5748..5770
 XX 5771..5793
 XX 5794..5816
 XX 5817..5839
 XX 5840..5862
 XX 5863..5885
 XX 5886..5908
 XX 5909..5931
 XX 5932..5954
 XX 5955..5977
 XX 5978..6000
 XX 6001..6023
 XX 6024..6046
 XX 6047..6069
 XX 6070..6092
 XX 6093..6115
 XX 6116..6138
 XX 6139..6161
 XX 6162..6184
 XX 6185..6207
 XX 6208..6230
 XX 6231..6253
 XX 6254..6276
 XX 6277..6299
 XX 6300..6322
 XX 6323..6345
 XX 6346..6368
 XX 6369..6391
 XX 6392..64

```
FT FT /label= X1
FT FT 372..392
FT FT /label= Y1
XX
PN DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 510; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
DB
RESULT 19
AAR66260
ID AAR66260 standard; protein; 392 AA.
XX
XX AAR66260;
XX
XX 25-MAR-2003 (updated)
XX 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M27.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Region 1..365
XX /label= M4
XX /note= "unglycosylated prourokinase (Ser47-Leu411)"
XX
XX Disulfide-bond 4..85
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX
```

```
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371
FT /label= X1
FT /label= Y1
FT
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX Whendt S;
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
XX improved fibrinolytic and thrombin inhibiting activities, for
XX treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
XX M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX are specific examples of such derivs. which have both improved
XX fibrinolytic and thrombin-inhibiting activities, compared to known
XX plasminogen activators or thrombin inhibitors. The proteins are
XX useful as thrombolytic agents, e.g. for treatment of arterial
XX occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX and pulmonary embolism.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 392 AA;
XX
XX Query Match 100.0%; Score 510; DB 16; Length 392;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLTQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
DB
RESULT 20
AAR66261
ID AAR66261 standard; protein; 392 AA.
XX
XX AAR66261;
XX
XX 25-MAR-2003 (updated)
XX 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M28.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX urokinase; variant; mutein.
XX
XX Synthetic.
XX
```

Key Location/Qualifiers
Region 1..365
/label= M4
/note= "unglycosylated prourokinase (Ser47-Leu411)"
XX Disulfide-bond 25..67
XX Disulfide-bond 56..80
XX Disulfide-bond 102..233
XX Disulfide-bond 143..159
XX Disulfide-bond 151..222
XX Disulfide-bond 247..316
XX Disulfide-bond 279..295
XX Disulfide-bond 306..334
Region 366..371
/label= X1
Region 372..392
/label= Y1
DE4323754-C1.
01-DEC-1994.
15-JUL-1993; 93DE-4323754.
15-JUL-1993; 93DE-4323754.
(CHEF) GRUENTHAL GMBH.
Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
Wendt S;
WPI; 1995-015191/03.
New bifunctional urokinase derivs and related plasmids - with improved fibrinolytic and thrombin inhibiting activities, for treating cardiac and cerebral infarct, pulmonary embolism, etc
Example 1; Page 11 and Fig 1; 34pp; German.
Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266 are specific examples of such derivs. which have both improved fibrinolytic and thrombin-inhibiting activities, compared to known plasminogen activators or thrombin inhibitors. The proteins are useful as thrombolytic agents, e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac and cerebral infarction and pulmonary embolism.
(Updated on 25-MAR-2003 to correct PN field.)
Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Y 1 KTCYEGNGHYRGKASDTMTGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDN 60
b 2 KTCYEGNGHYRGKASDTMTGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDN 61
Y 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
b 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 21
AR66263
D AAR66263 standard; protein; 392 AA.
X C AAR66263;
X 25-MAR-2003 (updated)
7-AUG-1995 (first entry)

DE Bifunctional urokinase variant M30.
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; muteln.
XX Synthetic.
XX Location/Qualifiers
Key 1..365
Region /label= M4
/note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
Region 366..371
/label= X1
Region 372..392
/label= Y1
DE4323754-C1.
01-DEC-1994.
15-JUL-1993; 93DE-4323754.
15-JUL-1993; 93DE-4323754.
(CHEF) GRUENTHAL GMBH.
Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
Wendt S;
WPI; 1995-015191/03.
New bifunctional urokinase derivs and related plasmids - with improved fibrinolytic and thrombin inhibiting activities, for treating cardiac and cerebral infarct, pulmonary embolism, etc
Example 1; Page 11 and Fig 1; 34pp; German.
Bifunctional urokinase derivatives corresponding to the formula M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266 are specific examples of such derivs. which have both improved fibrinolytic and thrombin-inhibiting activities, compared to known plasminogen activators or thrombin inhibitors. The proteins are useful as thrombolytic agents, e.g. for treatment of arterial occlusions, deep vein thrombosis, cardiac and cerebral infarction and pulmonary embolism.
(Updated on 25-MAR-2003 to correct PN field.)
Sequence 392 AA;
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHYRGKASDTMTGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHYRGKASDTMTGRPCLPWNSATVLOQTHAHRSDALQLGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 22
AAR66264


```
ID AAR66264 standard; protein; 392 AA.
XX AC AAR66264;
XX
XX 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M31.
DE
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..371 /label= X1
FT Region 372..392 /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 11 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence .392 AA;
XX
Query Match 100.0%; Score 510; DB 16; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.6e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 KTCYEGNGHGYRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHGYRGKASTDTWGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
```

```
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
RESULT 23
AAR66244
ID AAR66244 standard; protein; 393 AA.
XX
XX AAR66244;
XX
XX 25-MAR-2003 (updated)
DT 22-AUG-1995 (first entry)
XX
XX Bifunctional urokinase variant M11.
DE
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase(Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..374 /label= X1
FT Region 375..393 /label= Y1
XX
XX DE4323754-Cl.
XX
XX 01-DEC-1994.
XX
XX 15-JUL-1993; 93DE-4323754.
XX
XX 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX
XX WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence .393 AA;
XX
Query Match 100.0%; Score 510; DB 16; Length 393;
```

Best Local Similarity 100.0%; Pred. No. 4.7e-41; Indels 0; Gaps 0;
Matches 88; Conservative 0; Mismatches 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 24
AAR66249
ID AAR66249 standard; protein; 393 AA.
XX AC AAR66249;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M16.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..372 /label= X1
FT Region 373..393 /label= Y1
XX
PN DE4323754-Cl.
XX
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wndt S;
XX
DR WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
FT improved fibrinolytic and thrombin inhibiting activities, for
FT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial

CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 393 AA;
Query Match 100.0%; Score 510; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. No. 4.7e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 60
Db 2 KTCYEGNGHFYRGKASTDTMGRCPLPWSATVLIQOQTYHAHRSDALQGLGKHNYCRNPDN 61
QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 25
AAR66250
ID AAR66250 standard; protein; 393 AA.
XX AC AAR66250;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M17.
XX
XX fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365 /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Region 366..372 /label= X1
FT Region 373..393 /label= Y1
XX
PN DE4323754-Cl.
XX
XX
PD 01-DEC-1994.
XX
PF 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
XX (CHEF) GRUENENTHAL GMBH.
XX
XX Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wndt S;
XX
DR WPI; 1995-015191/03.
XX
XX New bifunctional urokinase derivs and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX Example 1; Page 10 and Fig 1; 34pp; German.
PS

XX Bifunctional urokinase derivatives corresponding to the formula
CC M4-XI-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 393 AA;

Query Match 100.0%; Score 510; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. NO. 4.7e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 26
AAR66251
ID AAR66251 standard; protein; 393 AA.
XX
AC AAR66251;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M18.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region /label= X1
FT Region /label= Y1
XX
PN DE4323754-C1.
XX
PD 01-DEC-1994.
XX
PP 15-JUL-1993; 93DE-4323754.
XX
PR 15-JUL-1993; 93DE-4323754.
XX
PA (CHEF) GRUENENTHAL GMBH.
XX
PI Heinzel-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
PI Wnendt S;
XX

DR WPI; 1995-015191/03.
XX
PT New bifunctional urokinase derive and related plasmids - with
PT improved fibrinolytic and thrombin inhibiting activities, for
PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
CC Bifunctional urokinase derivatives corresponding to the formula
CC M4-XI-Y1 are claimed (see features table). Sequences AAR66244-R66266
CC are specific examples of such derivs. which have both improved
CC fibrinolytic and thrombin-inhibiting activities, compared to known
CC plasminogen activators or thrombin inhibitors. The proteins are
CC useful as thrombolytic agents, e.g. for treatment of arterial
CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
CC and pulmonary embolism.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 393 AA;

Query Match 100.0%; Score 510; DB 16; Length 393;
Best Local Similarity 100.0%; Pred. NO. 4.7e-41;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 60
DB 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLQOTYHAHRSALQLGLGKHNYCRNPDN 61

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
DB 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89

RESULT 27
AAR66252
ID AAR66252 standard; protein; 393 AA.
XX
AC AAR66252;
XX
DT 25-MAR-2003 (updated)
DT 17-AUG-1995 (first entry)
XX
DE Bifunctional urokinase variant M19.
XX
KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
KW urokinase; variant; mutein.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..365
FT /label= M4
FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
FT Disulfide-bond 4..85
FT Disulfide-bond 25..67
FT Disulfide-bond 56..80
FT Disulfide-bond 102..233
FT Disulfide-bond 143..159
FT Disulfide-bond 151..222
FT Disulfide-bond 247..316
FT Disulfide-bond 279..295
FT Disulfide-bond 306..334
FT Disulfide-bond 366..372
FT Region /label= X1
FT Region /label= Y1
XX
PN DE4323754-C1.
XX
PD 01-DEC-1994.
XX
PP 15-JUL-1993; 93DE-4323754.
XX

```
PR 15-JUL-1993; 93DE-4323754.
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wndt S;
XX DR WPI; 1995-015191/03.
XX
XX PT New bifunctional urokinase derivs and related plasmids - with
XX PT improved fibrinolytic and thrombin inhibiting activities, etc
XX PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 393 AA;
XX
XX Query Match 100.0%; Score 510; DB 16; Length 393;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
XX
XX Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
XX 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
XX
XX RESULT 28
XX AAR66253
XX ID AAR66253 standard; protein; 393 AA.
XX
XX AC AAR66253;
XX
XX XX 25-MAR-2003 (updated)
XX DT 17-AUG-1995 (first entry)
XX
XX DE Bifunctional urokinase variant M20.
XX
XX KW fibrinolysis; thrombin inhibition; thrombolytic; anti-thrombotic;
XX KW urokinase; variant; mutein.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Region /label= M4
XX FT /note= "unglycosylated prourokinase (Ser47-Leu411)"
XX FT 4..85
XX FT Disulfide-bond 25..67
XX FT Disulfide-bond 56..80
XX FT Disulfide-bond 102..233
XX FT Disulfide-bond 143..159
XX FT Disulfide-bond 151..222
XX FT Disulfide-bond 247..316
XX FT Disulfide-bond 279..295
XX FT Disulfide-bond 306..334
XX FT Region 366..372
XX FT /label= X1
XX FT 373..393
XX FT Region /label= Y1
XX FT
```

```
XX DE4323754-C1.
XX
XX PD 01-DEC-1994.
XX
XX PF 15-JUL-1993; 93DE-4323754.
XX
XX PR 15-JUL-1993; 93DE-4323754.
XX
XX PA (CHEF ) GRUENENTHAL GMBH.
XX
XX PI Heinzl-Wieland R, Saunders DJ, Schneider J, Steffens GJ;
XX PI Wndt S;
XX
XX DR WPI; 1995-015191/03.
XX
XX PT New bifunctional urokinase derivs and related plasmids - with
XX PT improved fibrinolytic and thrombin inhibiting activities, for
XX PT treating cardiac and cerebral infarct, pulmonary embolism, etc
XX
XX PS Example 1; Page 10 and Fig 1; 34pp; German.
XX
XX CC Bifunctional urokinase derivatives corresponding to the formula
XX CC M4-X1-Y1 are claimed (see features table). Sequences AAR66244-R66266
XX CC are specific examples of such derivs. which have both improved
XX CC fibrinolytic and thrombin-inhibiting activities, compared to known
XX CC plasminogen activators or thrombin inhibitors. The proteins are
XX CC useful as thrombolytic agents, e.g. for treatment of arterial
XX CC occlusions, deep vein thrombosis, cardiac and cerebral infarction
XX CC and pulmonary embolism.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 393 AA;
XX
XX Query Match 100.0%; Score 510; DB 16; Length 393;
XX Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 60
XX 2 KTCYEGNGHFYRGKASTDTMGRPCLPWNSATVLOQTYHAHRSDALQLGLGKHNYCRNPDN 61
XX
XX Db 61 RRRPWCYVQVGLKPLVQECMVHDCADGK 88
XX 62 RRRPWCYVQVGLKPLVQECMVHDCADGK 89
XX
XX RESULT 29
XX AAR99885
XX ID AAR99885 standard; peptide; 393 AA.
XX
XX AC AAR99885;
XX
XX DT 27-JAN-1997 (first entry)
XX
XX DE M36: fibrinolytic and anticoagulant activity contg. protein.
XX
XX KW Thrombin; factor Xa; anticoagulant; thrombolytic; plasminogen;
XX KW urokinase; activator; streptokinase; staphylokinase; APSAC;
XX KW anisolated plasminogen streptokinase activator complex; hirudin;
XX KW hirullin; antistatin; pWLT27; pWS1; pSE8; pW56.
XX
XX OS Synthetic.
XX
XX PN EP712934-A2.
XX
XX PD 22-MAY-1996.
XX
XX PF 03-NOV-1995; 95EP-0117316.
XX
XX PR 17-NOV-1994; 94DE-4440892.
XX
XX PA (CHEF ) GRUENENTHAL GMBH.
```

```

XX Heinzel-Wieland R, Steffens GJ, Wnendt S;
XX WPI; 1996-240720/25.
XX
XX Proteins with fibrinolytic and anticoagulant activity - useful as
XX thrombolytic agents
XX
XX Disclosure; Fig 18; 59pp; German.
XX
XX New peptides (I) with fibrinolytic and anticoagulant activity
XX comprise a plasminogen-activating amino acid sequence (A) fused
XX at the N- and/or C-terminus to a thrombin and/or factor Xa
XX inhibiting amino acid sequence (B). Excluded from the claims
XX are (I) where (A) is Ser47 to Leu411 of unglycosylated urokinase
XX linked at the C-terminus to sequences (i) to (iii):
XX   TI-RP-T2-GGGGNGDFEIEPEYL-T3      (i)
XX   TI-RPFLRNPNKDYEPFWEDEKNE        (ii)
XX   TI-RPSSEPEFEIDEEK                (iii)
XX   Where T1= P or V; T2 = L or a bond; T3= Q or OH.
XX   (A) is pref. (pro)urokinase; tissue plasminogen activator (tPA),
XX   bat-PA (all opt. modified by deletion, substitution, insertion and/or
XX   addn.); streptokinase; staphylokinase; and/or APSAC (anisolated
XX   plasminogen streptokinase activator complex), esp. prourokinase
XX   (411 amino acids) or its Ser47 to Leu411 or Ser138 to Leu411
XX   fragments, or t-PA (527 amino acids) or its Ser89Arg to 527Pro or
XX   174Ser to 527Pro fragments.
XX   (B) has hirudin or hirullin activities; or is derived from the human
XX   thrombin receptor, anisatin and/or the tick anticoagulant peptide.
XX   Most pref. are the 65 amino acid hirudin sequence or one of the six
XX   sequences given in AAR99879 to AAR99884.
XX   Plasmids pMT127 (M51), pWS1 (M5112), pSE8 (M36) and pW56 (M43)
XX   contain the sequences encoding AAR99885 to AAR99889, respectively.
XX   The products were tested in human citrated plasma (5 microg in 200
XX   microl 1:10 diluted plasma). The thrombin time was then 1.2, 3, 2.8
XX   and 1.2 times greater, respectively, than in the absence of the product.
XX
XX   Sequence 393 AA;
XX
XX   Query Match 100.0%; Score 510; DB 17; Length 393;
XX   Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX   Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX   QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
XX   Db 3 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 62
XX
XX   QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
XX   Db 63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90
XX
XX   Search completed; December 3, 2003, 14:39:09
XX   Job time : 14.8621 secs
XX
XX   Query Match 100.0%; Score 510; DB 17; Length 393;
XX   Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX   Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX   QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
XX   Db 3 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 62
XX
XX   QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
XX   Db 63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90
XX
XX   RESULT 30
XX   AAR99596
XX   ID AAR99596 standard; Protein; 393 AA.
XX   AC AAR99596;
XX   DT 05-DEC-1996 (first entry)
XX   DE Chimeric protein M37 encoded by p859.
XX   KW Thrombin; inhibition; thrombus; thrombolysis; chimeric protein;
XX   KW plasminogen activating sequence; fibrinolysis; infarction;
XX   KW angina pectoris; deep vein thrombosis.
XX   OS Synthetic.
XX   PN EF714982-A2.
XX   PD 05-JUN-1996.
XX   PF 16-NOV-1995; 95EP-0118050.

```

```

XX 30-NOV-1994; 94DE-4442665.
XX (CHEF ) GRUENENTHAL GMEH.
XX Heinzel-Wieland R, Janocha E, Steffens GJ, Wnendt S;
XX WPI; 1996-269715/28.
XX
XX Chimeric protein contg. plasminogen activating sequence and
XX thrombin-inhibiting sequence - useful as thrombus-specific
XX thrombolytic agent with rapid action
XX
XX Example 1; Page 19-20; 37pp; German.
XX
XX Example 1 describes the prodn. of plasmids pSE1 and pSE9
XX contg. a DNA encoding a chimeric protein with fibrinolytic and
XX thrombin-inhibiting properties.
XX pSE1 encodes the protein given in AAR99597 and pSE9 encodes
XX the protein given in AAR99596.
XX
XX   Sequence 393 AA;
XX
XX   Query Match 100.0%; Score 510; DB 17; Length 393;
XX   Best Local Similarity 100.0%; Pred. No. 4.7e-41;
XX   Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX   QY 1 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
XX   Db 3 KTCYEGNGHFYRGKASTDTWGRPCLPWNSATVLQOTYHAHRSDALQLGLGKHNCRNPDN 62
XX
XX   QY 61 RRRPWCYVQVGLKPLVQECWVHDCADGK 88
XX   Db 63 RRRPWCYVQVGLKPLVQECWVHDCADGK 90
XX
XX   Search completed; December 3, 2003, 14:39:09
XX   Job time : 14.8621 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 3, 2003, 14:34:23 ; Search time 10.8276 Seconds
(without alignments)
2097.294 Million cell updates/sec

Title: US-09-880-503-1
Perfect score: 510
Sequence: 1 KTCVEGNHFRGRKASTDM.....QVGLKPLVQECMWHDCADGK 88

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	495	97.1	154	4 Q96SE8	Q96SE8 homo sapien
2	415	81.4	433	6 Q8MIL0	Q8MIL0 oryctolagus
3	415	81.4	433	6 Q8MHY7	Q8MHY7 oryctolagus
4	396	77.6	157	6 Q9TVA8	Q9TVA8 bos taurus
5	387	75.9	231	11 Q8C6L2	Q8C6L2 mus musculus
6	313	61.4	214	6 Q9XT70	Q9XT70 oryctolagus
7	235	46.1	564	6 Q8KEB1	Q8KEB1 oryctolagus
8	226	44.3	516	4 Q9BU99	Q9BU99 homo sapien
9	219	42.9	559	11 Q91VP2	Q91VP2 mus musculus
10	218.5	42.8	90	4 Q8NG20	Q8NG20 homo sapien
11	215	42.2	395	4 Q9BZW1	Q9BZW1 homo sapien
12	210	41.2	562	6 Q8SQ23	Q8SQ23 mus musculus
13	209.5	41.1	517	11 Q8KOD2	Q8KOD2 mus musculus
14	193.5	37.9	560	4 Q14520	Q14520 homo sapien
15	191	37.5	202	13 Q90675	Q90675 gallus gall
16	188	36.9	653	11 Q8VCS4	Q8VCS4 mus musculus

17	186.5	36.6	597	11 Q35727	Q35727 mus musculus
18	177.5	34.8	616	6 Q97507	Q97507 sus scrofa
19	172.5	33.8	421	13 Q8AXX3	Q8AXX3 xenopus lae
20	170.5	33.4	615	4 Q81ZZ5	Q81ZZ5 homo sapien
21	163.5	32.1	947	13 Q8AXY6	Q8AXY6 gallus gall
22	158	31.0	452	13 Q90Y90	Q90Y90 xenopus lae
23	158	31.0	454	6 Q46506	Q46506 papio hamad
24	156	30.6	113	4 Q9UIR5	Q9UIR5 homo sapien
25	154	30.2	806	6 Q18783	Q18783 macropus eu
26	153	30.0	113	4 Q9UIR7	Q9UIR7 homo sapien
27	151.5	29.7	359	6 Q8WMR1	Q8WMR1 canis famil
28	149.5	29.3	812	11 Q9R0W3	Q9R0W3 rattus norv
29	149	29.2	716	13 Q91691	Q91691 xenopus lae
30	148.5	29.1	381	4 Q8N2U4	Q8N2U4 homo sapien
31	148.5	29.1	399	4 Q96GL8	Q96GL8 homo sapien
32	148.5	29.1	420	4 Q9BTP9	Q9BTP9 homo sapien
33	148.5	29.1	424	4 Q8NCW1	Q8NCW1 homo sapien
34	148.5	29.1	462	4 Q8NCW0	Q8NCW0 homo sapien
35	148.5	29.1	801	11 Q8KQ08	Q8KQ08 mus musculus
36	148.5	29.1	944	11 Q8C3W2	Q8C3W2 mus musculus
37	148.5	29.1	944	11 Q8BSP6	Q8BSP6 mus musculus
38	148	29.0	105	4 Q9UIR8	Q9UIR8 homo sapien
39	147	28.8	113	4 Q9UIR6	Q9UIR6 homo sapien
40	147	28.8	648	4 Q9H1V4	Q9H1V4 homo sapien
41	146.5	28.7	461	11 Q8KIS7	Q8KIS7 mus musculus
42	146	28.6	393	4 Q9BRB6	Q9BRB6 homo sapien
43	146	28.6	567	4 Q13208	Q13208 homo sapien
44	146	28.6	937	11 Q8BNP9	Q8BNP9 mus musculus
45	146	28.6	937	11 Q8BG10	Q8BG10 mus musculus

ALIGNMENTS

RESULT 1

Q96SE8 ID Q96SE8 PRELIMINARY; PRT; 154 AA.

AC Q96SE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator amino-terminal fragment.
GN ATF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J., Bai X., Ruan C.;
RT "Cloning and expression of the amino-terminal fragment of human urokinase-type plasminogen activator."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Bai X., Fu J., Wang W., Xi X., Ruan C.;
RT "Overexpression of the amino-terminal fragment of human urokinase-type plasminogen activator in breast cancer cells results in decreased tumor invasion, growth and angiogenesis."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029537; AAK38734.1; -.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle_1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle.
DR SMART; SM00130; KR; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.

```
KW Glycoprotein; Kinase; Kringle.
SQ SEQUENCE 154 AA; 17305 MW; A3CCF2FCFF505572 CRC64;

Query Match          97.1%; Score 495; DB 4; Length 154;
Best Local Similarity 98.9%; Pred. No. 1.6e-51;
Matches 86; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
68 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 127
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 87
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 RRRPWCYVQVGLKPLVQECMVHDCADG 154
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 2
Q8MHY7
ID Q8MHY7 PRELIMINARY; PRT; 433 AA.
AC Q8MHY7;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase-type plasminogen activator.
GN UROKINASE.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yano W., Watanabe M.;
RL "Oryctolagus cuniculus urokinase-type plasminogen activator, mRNA,
RT complete cds.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AY029517; AAK40239.1; -.
DR EMBL; AB087224; BAC02685.1; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Serprotease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS50240; TRYPSIN_HIS; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
SQ SEQUENCE 433 AA; 48444 MW; 6DD35A371010A6EE CRC64;

Query Match          81.4%; Score 415; DB 6; Length 433;
Best Local Similarity 79.5%; Pred. No. 2e-41;
Matches 70; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 1 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
70 KTCYEGNGHFYRGKASTDTMGPRCLPWN SATVLQOTYHAHRSDALQLGLGKHNCRNPDN 129
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 RRRPWCYVQVGLKPLVQECMVHDCADG 88
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 130 QRRPWCYVQVGLKPLVQECMVHDCSSGK 157
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
Q9TVAS
ID Q9TVAS PRELIMINARY; PRT; 157 AA.
AC Q9TVAS;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Urokinase plasminogen activator (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
```

Db 129 QXRPCYVQIGLRQFVQECNVHDCSLSK 156

RESULT 6
Q9XT70

ID Q9XT70 PRELIMINARY; PRT; 214 AA.

AC Q9XT70;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urokinase-type plasminogen activator (Fragment).

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Lung;

RA "Partial mRNA of rabbit uPA."
RA Yin J., Idell S.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -! SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF097647; AAD39351.1; --
DR HSSP; P00749; 1EJN.
DR MEROPS; S01.231; --
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; KcKring; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; TRY_SPc; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYP SIN_HIS; 1.
DR PROSITE; PS00134; TRYP SIN_DOM; 1.
KW Glycoprotein; Hydrolase; Kinase; Kringle; Protease; Serine protease.
FT NON_TER 214 214
SQ SEQUENCE 214 AA; 24314 MW; 69975C41C32B0D7E CRC64;

Query Match 61.4%; Score 313; DB 6; Length 214;
Best Local Similarity 78.3%; Pred. No. 1.6e-29;
Matches 54; Conservative 5; Mismatches 10; Indels 0; Gaps 0

Qy 20 MGRCLPNWATVLQOTYHAHRSDALQLGLGHNYCRNPDPNRPPWCYVQVGLKPVLQEC 79
Db 1 MDRECLAWNSANLVTKTYHAHPDALQLGLGHNYCRNPDPHORPPWCYVQVGLKQLIQEC 60
Qy 80 MVHDCADGK 88
Db 61 KVHDSSGGK 69

RESULT 7
Q8MKB1

ID Q8MKB1 PRELIMINARY; PRT; 564 AA.

AC Q8MKB1;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Tissue-type plasminogen activator.

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;

RN [1]

RP SEQUENCE FROM N.A.

RA Sugiki M., Yoshida E., Anai K., Maruyama M.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	EMBL; AY029518; AAK40240.1; -.	
DR	InterPro; IPR001314; Chymotrypsin.	
DR	InterPro; IPR006209; EGF like.	
DR	InterPro; IPR000083; Fibnctnl.	
DR	InterPro; IPR006210; IEFG.	
DR	InterPro; IPR000001; Kringle.	
DR	InterPro; IPR001254; Ser_protease_Try.	
DR	Pfam; PF00008; EGF; 1.	
DR	Pfam; PF00039; fnl; 1.	
DR	Pfam; PF00051; kringle; 2.	
DR	Pfam; PF00089; trypsin; 1.	
DR	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	PRINTS; PR00018; KRINGLE.	
DR	ProDom; PD000395; Kringle; 2.	
DR	SMART; SM00181; EGF; 1.	
DR	SMART; SM00130; KR; 2.	
DR	PROSITE; PS00020; Tryp_Spc; 1.	
DR	PROSITE; PS00022; EGF_1; 1.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS00021; KRINGLE_1; 2.	
DR	PROSITE; PS00070; KRINGLE_2; 2.	
DR	PROSITE; PS02040; TRYPSIN_DOM; 1.	
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.	
DR	PROSITE; PS00135; TRYPSIN_SER; 1.	
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;	
KW	Serine protease.	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 564 AA; 62726_MW; 459D8BAC6D4A937C CRC64;	
QY	Query Match	46.1%; Score 235; DB 6; Length 564;
QY	Best Local Similarity	50.6%; Pred. No. 1.le-19;
QY	Matches	42; Conservative 10; Mismatches 31; Indels 0; Gaps 0;
QY	3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 62	
Db	217 CYLKGQAYRGTHSLTTTSGASCPLWNSMLLVGEKTYARQSNABALGLGKHNYCRNPDGDS 276	
QY	63 RRPWCYQVQGLKPLVQECMVHDCA 85	
Db	277 KPWCHLVKNKRLVIEYCDVPOCA 299	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	

CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.	
CC	EMBL; AY029518; AAK40240.1; -.	
DR	InterPro; IPR001314; Chymotrypsin.	
DR	InterPro; IPR006209; EGF like.	
DR	InterPro; IPR000083; Fibnctnl.	
DR	InterPro; IPR006210; IEFG.	
DR	InterPro; IPR000001; Kringle.	
DR	InterPro; IPR001254; Ser_protease_Try.	
DR	Pfam; PF00008; EGF; 1.	
DR	Pfam; PF00039; fnl; 1.	
DR	Pfam; PF00051; kringle; 2.	
DR	Pfam; PF00089; trypsin; 1.	
DR	PRINTS; PR00722; CHYMOTRYPSIN.	
DR	PRINTS; PR00018; KRINGLE.	
DR	ProDom; PD000395; Kringle; 2.	
DR	SMART; SM00181; EGF; 1.	
DR	SMART; SM00130; KR; 2.	
DR	PROSITE; PS00020; Tryp_Spc; 1.	
DR	PROSITE; PS00022; EGF_1; 1.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS00021; KRINGLE_1; 2.	
DR	PROSITE; PS00070; KRINGLE_2; 2.	
DR	PROSITE; PS02040; TRYPSIN_DOM; 1.	
DR	PROSITE; PS00134; TRYPSIN_HIS; 1.	
DR	PROSITE; PS00135; TRYPSIN_SER; 1.	
KW	EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;	
KW	Serine protease.	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 564 AA; 62726_MW; 459D8BAC6D4A937C CRC64;	
QY	Query Match	46.1%; Score 235; DB 6; Length 564;
QY	Best Local Similarity	50.6%; Pred. No. 1.le-19;
QY	Matches	42; Conservative 10; Mismatches 31; Indels 0; Gaps 0;
QY	3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 62	
Db	217 CYLKGQAYRGTHSLTTTSGASCPLWNSMLLVGEKTYARQSNABALGLGKHNYCRNPDGDS 276	
QY	63 RRPWCYQVQGLKPLVQECMVHDCA 85	
Db	277 KPWCHLVKNKRLVIEYCDVPOCA 299	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	
Db	140 SKPCYVFKAGKYSSEFCSTPAASEG 165	
QY	SEQUENCE 516 AA; 57370_MW; BAB31901FDC96800 CRC64;	
QY	Query Match	44.3%; Score 226; DB 4; Length 516;
QY	Best Local Similarity	47.7%; Pred. No. 1.le-18;
QY	Matches	41; Conservative 9; Mismatches 36; Indels 0; Gaps 0
QY	2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQTYHAHRSALQLGLGKHNCVCRPDNR 61	
Db	80 TCYEDQGISYRGTWSTAESGAECTNWNSSALQAQPSYGRPRPDATRLGLGNHNYCRNPDNR 139	
QY	62 RRPWCYQVQGLKPLVQECMVHDCADG 87	

```

DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 2.
DR PROSITE; PS50070; KRINGLE_2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 559 AA; 63122 MW; 8CCE2BDB94514D9 CRC64;

Query Match 42.9%; Score 219; DB 11; Length 559;
Best Local Similarity 46.0%; Pred. No. 8.5e-18;
Matches 40; Conservative 11; Mismatches 36; Indels 0; Gaps 0;

QY 2 TCYEGNGHFYRGKASDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPDNR 61
Db 123 TCYEGQITVYRGWTSABSGAECINWNSVLSLKPNARRPNAIKLGLGNHNYCRNPDNR 182

QY 62 RRPWCYVQVGLKPLVQECMVHDCAGK 88
Db 183 LKPCWYVFKAGYTTTFCSPFACPKGK 209

RESULT 10
Q8NG20 PRELIMINARY; PRT; 90 AA.
AC Q8NG20;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Plasminogen/activator kringle.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

Query Match 42.8%; Score 218.5; DB 4; Length 90;
Best Local Similarity 49.4%; Pred. No. 1.3e-18;
Matches 41; Conservative 5; Mismatches 36; Indels 1; Gaps 1;

QY 3 CYEGNGHFYRGKASDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPDNR 62
Db 8 CYFNGGAYRGTHSLTESGASCLPWN SMILIGKYVTAQNPSAQLGLGKHN YCRNPDGDA 67

QY 63 RWCYVQVGLKPLVQECMVHDC 85
Db 68 KPCWYV-TNPRKLYDCVDPQCA 89

RESULT 11
Q9BZW1 PRELIMINARY; PRT; 395 AA.
AC Q9BZW1;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Neonatal thrombolytic agent alpha-form (Fragment).

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Dou D.;
RT "A brain-type plasminogen activator.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF260825; AAK11956.1; -.
DR HSSP; P00750; 1PK2.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00039; fnl; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00038; FNI; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00020; TRYPSIN; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 395
SQ SEQUENCE 395 AA; 44323 MW; 3FBD4A2F0B7C11C8 CRC64;

Query Match 42.2%; Score 215; DB 4; Length 395;
Best Local Similarity 47.0%; Pred. No. 1.8e-17;
Matches 39; Conservative 8; Mismatches 36; Indels 0; Gaps 0;

QY 3 CYEGNGHFYRGKASDTMGRCPLPWN SATVLQOTYHAHRS DALQLGLGKHN YCRNPDNR 62
Db 48 CYFNGGAYRGTHSLTESGASCLPWN SMILIGKYVTAQNPSAQLGLGKHN YCRNPDGDA 107

QY 63 RWCYVQVGLKPLVQECMVHDC 85
Db 108 KPCWCHLVKRNRLTWECYDVPSCS 130

RESULT 12
Q8SQ23 PRELIMINARY; PRT; 562 AA.
AC Q8SQ23;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE T-plasminogen activator.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Enamel Organ;
RA Ding Y., Xue J., Bartlett J.D.;
RT "T-plasminogen activator in tooth tissues.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; AF364605; AAM00297.1; -.
DR HSSP; P00761; 1ANI.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctnl.
DR InterPro; IPR006210; IEGF.

```

```
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00039; fnl; 1.
DR Pfam: PF00051; kringle; 2.
DR Pfam: PF00089; trypsin; 1.
DR ProDom: PD000395; Kringle; 2.
DR SMART: SM00181; EGF; 1.
DR SMART: SM00058; fnl; 1.
DR SMART: SM00130; KR; 2.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBONECTIN_1; 1.
DR PROSITE: PS00021; KRINGLE_1; 2.
DR PROSITE: PS00070; KRINGLE_2; 2.
DR PROSITE: PS02040; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Protease;
SQ SEQUENCE 562 AA; 63668 MW; F9E6B4C77CB101E8 CRC64;

Query Match 41.2%; Score 210; DB 6; Length 562;
Best Local Similarity 43.7%; Pred. No. 1e-16;
Matches 38; Conservative 9; Mismatches 40; Indels 0; Gaps 0;

QY 2 TCYEGNGHYRGKASTDTMGRPCLPNWSATVLQOTYHAHRSALQLGLGKHNYCRNPDNR 61
Db 126 TCYEGDGIYRGTWSTESGAECVNNWNTSLASMPYNGRRPDVAKVLGLGNHNYCRNPKD 185
QY 62 RRPWCYVQVGLKPLVQECMVHDCADGK 88
Db 186 SRPWCYIFKAEKXSPDFCSTPACTKEK 212

RESULT 13
Q8K0D2 PRELIMINARY; PRT; 517 AA.
AC Q8K0D2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
CC -!- SIMILARITY: BELONGS TO THE EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC031775; AAH31775.1; -.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR006209; EGF-like.
DR InterPro: IPR006210; IEGF.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00008; EGF; 2.
DR Pfam: PF00051; kringle; 1.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00022; EGF_1; 3.
```

```
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 517 AA; 57326 MW; 3855A42035A5EA59 CRC64;

Query Match 41.1%; Score 209.5; DB 11; Length 517;
Best Local Similarity 45.8%; Pred. No. 1.1e-16;
Matches 38; Conservative 13; Mismatches 31; Indels 1; Gaps 1;

QY 3 CYEGNGHYRGKASTDTMGRPCLPNWSATVLQOTYHAHRSALQLGLGKHNYCRNPDNR 62
Db 151 CYVGDGYSYRGKSVKTVNQNPCLYWNHLLQETYNFMEDAETHGIAEHFRCRNPDDGH 210
QY 63 RWCYVQVGLKPLVQECMVHDC 84
Db 211 KPCFCYKVNSEKVKMEYCDTVTC 233

RESULT 14
Q14520 PRELIMINARY; PRT; 560 AA.
AC Q14520; O00663;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HGF activator like protein (Hyaluronan binding protein 2).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RA Kitamura N.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [2];
RP SEQUENCE FROM N.A.
RX MEDLINE=96425001; PubMed=8827452;
RA Choi-Mura N.H., Tobe T., Sumiya J., Nakano Y., Sano Y., Mazda T.,
RA Tomita M.;
RT "Purification and characterization of a novel hyaluronan-binding
RT protein (PHBP) from human plasma: it has three EGF, a kringle and a
RT serine protease domain, similar to hepatocyte growth factor
RT activator.";
RL J. Biochem. 119:1157-1165(1996).
RN [3];
RP SEQUENCE FROM N.A.
RC TISSUE=Colon, and Kidney;
RA Strausberg R.;
CC -!- SIMILARITY: BELONGS TO THE EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; D49742; BAA08576.1; -.
DR EMBL; S83182; AAB46909.1; -.
DR EMBL; BC031412; AAH31412.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.033; -.
DR Genew; HGNC:4798; HABP2.
DR InterPro: IPR001314; Chymotrypsin.
DR InterPro: IPR006209; EGF like.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR ProDom; PD000395; Kringle.
DR PROSITE; PS00022; EGF_1; 3.
```

```
DR PROSITE; PS01186; EGF 2; 2.
DR PROSITE; PS00021; KRINGLE 1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 560 AA; 62671 MW; 5C1907230784ACD4 CRC64;

Query Match 37.9%; Score 193.5; DB 4; Length 560;
Best Local Similarity 41.7%; Pred. No. 9.8e-15;
Matches 35; Conservative 15; Mismatches 33; Indels 1; Gaps 1;

QY 3 CYEENGHFYRGKASTDTMGRCPLPWN SATVLQQTYYHAHRS DALQLGLGKHN YCRNP DNR 62
Db 194 CYVGGSYRGKNNVTNQHACLYN SHLLIQENYMWEDAE THGIGEHFCRNPD ADE 253

QY 63 RPWCYVQVGLKPLVQECMVHDC 85
Db 254 KPWCIFKVTNDKVKWEYCDVSACS 277

RESULT 15
Q90675 PRELIMINARY; PRT; 202 AA.
ID AC Q90675;
DC 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Tissue-type plasminogen activator (Fragment).
GN TPA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC STRAIN=White Leghorn;
RX MEDLINE=97199025; PubMed=9047000;
RT Johnson A.L., Bridgham J.T., Anthony R.V.;
RT "Expression of avian urokinase and tissue-type plasminogen activator
RT messenger ribonucleic acid during follicle development and atresia.";
RL Biol. Record. 56:581-588(1997)
CC !- SIMILARITY: CONTAINS 2 KRINGLE DOMAINS.
DR EMBL; U31988; AAA74955.1; -.
DR HSSP; P00750; IRTF.
DR MEROPS; S01.232; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00051; kringle; 2.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 2.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_2.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE 2; 2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 202
SQ SEQUENCE 202 AA; 23148 MW; 049DB42941D83AFC CRC64;

Query Match 37.5%; Score 191; DB 13; Length 202;
Best Local Similarity 43.9%; Pred. No. 6.4e-15;
Matches 36; Conservative 7; Mismatches 39; Indels 0; Gaps 0;

QY 3 CYEENGHFYRGKASTDTMGRCPLPWN SATVLQQTYYHAHRS DALQLGLGKHN YCRNP DNR 62
Db 41 CYTGNGLAYRGTSRTKSGFCLPWN PVFLTSKIYTALEQRRA LGLGKHNCRNP DGA 100
```

```
QY 63 RPWCYVQVGLKPLVQECMVHDC 84
Db 101 QPWCYVWKDRQLTWYCDVPOC 122

RESULT 16
Q8VCS4 PRELIMINARY; PRT; 653 AA.
ID AC Q8VCS4;
DC 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 70.6 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC !- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC !- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; BC019376; AAH19376.1; -.
DR HSSP; P00761; IAN1.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser. protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hypothetical protein; EGF-like domain; Glycoprotein; Hydrolase;
KW Kringle; Protease; Serine protease.
SQ SEQUENCE 653 AA; 70553 MW; FE18D90174ED6FDD CRC64;

Query Match 36.9%; Score 188; DB 11; Length 653;
Best Local Similarity 53.0%; Pred. No. 5.3e-14;
Matches 35; Conservative 6; Mismatches 25; Indels 0; Gaps 0;

QY 3 CYEENGHFYRGKASTDTMGRCPLPWN SATVLQQTYYHAHRS DALQLGLGKHN YCRNP DNR 62
Db 283 CFLNGTEYRGVASTAASGLSLAWNSDLLYQELHVDVAAA VLLGLGPHAYCRNP DKE 342

QY 63 RPWCYV 68
Db 343 RPWCYV 348
```

RESULT 17

```
O35727 ID O35727 PRELIMINARY; PRT; 597 AA.
AC O35727 AC O35727 PRELIMINARY; PRT; 597 AA.
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Factor XII.
GN F12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RT Takahashi T., Kihara T.;
RA Schlosser M., Schwager S., Engel W.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; X95711; CAA67891.1; -.
DR HSSP; P00760; 1A07.
DR MEROPS; S01.211; -.
DR MGD; MGI:1891012; F12.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 597 AA; 65638 MW; F3AC07C37D0C0FEA CRC64;
```

```
Query Match 36.6%; Score 186.5; DB 11; Length 597;
Best Local Similarity 44.0%; Pred. No. 7.2e-14;
Matches 37; Conservative 8; Mismatches 34; Indels 5; Gaps 2;

QY 2 TCYEGNGHFYRGKASTDTMGPCLPWNSATVLQOTY-HAHRSDALQLGLGKHNYCRNPDN 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 TCYEGRLSYRGAGTQTSGAPCQRW---TVEATYRNMTKQALSWGLGHAFRCRPN 271
QY 61 RRPWCYVQVGLKPLVQECMVHDC 84
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 DTRPWCYVQVGLKPLVQECMVHDC 295
```

RESULT 18

```
O97507 ID O97507 PRELIMINARY; PRT; 616 AA.
AC O97507 AC O97507 PRELIMINARY; PRT; 616 AA.
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE FXII.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RT Takahashi T., Kihara T.;
RA "Porcine liver factor XII.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AB022426; BAA37148.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.211; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR006209; EGF_like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00039; fn1; 1.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000995; FN_Type_II; 1.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00181; EGF; 2.
DR SMART; SM00058; FN1; 1.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00130; KR; 1.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW EGF-like domain; Glycoprotein; Hydrolase; Kringle; Protease;
KW Serine protease.
SQ SEQUENCE 616 AA; 405FE3D71EBBD1A9 CRC64;
```

```
Query Match 34.8%; Score 177.5; DB 6; Length 616;
Best Local Similarity 44.6%; Pred. No. 9e-13;
Matches 37; Conservative 6; Mismatches 35; Indels 5; Gaps 2;

QY 3 CYEGNGHFYRGKASTDTMGPCLPWNSATVLQOTY-HAHRSDALQLGLGKHNYCRNPDN 61
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 CYSDRGLSYRGAGTQTLSGAPCQFAS---EATYRNMTAEQALNWGLGDHAFRCRPN 272
QY 62 RRPWCYVQVGLKPLVQECMVHDC 84
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 TRPWCYVQVGLKPLVQECMVHDC 295
```

RESULT 19

```

QBAXX3
ID QBAXX3 PRELIMINARY; PRT; 421 AA.
AC QBAXX3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Kremen2.
GN KRM2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22308873; PubMed=12421700;
RA Davidson G., Mao B., Del Barco Barrantes I., Niehrs C.;
RT "kremen proteins interact with Dickkopf1 to regulate anteroposterior
RL CNS patterning.";
RL Development 129:5587-5596(2002).
DR EMBL; AY150813; AAN64661.1; -.
SQ SEQUENCE 421 AA; 46274 MW; 74D04DB1682CD15F CRC64;

Query Match 33.8%; Score 172.5; DB 13; Length 421;
Best Local Similarity 50.7%; Pred. No. 2.4e-12;
Matches 34; Conservative 4; Mismatches 26; Indels 3; Gaps 2;

QY 3 CYEGNGHFYRGKAS-TDTMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRNPDR 61
DB 29 CTFVNGRDYRGTSVQAGPGTFLYWNQTT--QHLVNAQSDPDGELGLGNHNYCRNPDA 86
QY 62 RRPWCYV 68
DB 87 VDPWCYV 93

RESULT 20
QB1ZZ5
ID QB1ZZ5 PRELIMINARY; PRT; 615 AA.
AC QB1ZZ5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Coagulation factor XII-Mie.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Buthera; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wada H., Nishioka J., Nakatani K., Kasai Y., Abe Y., Nobori T.;
RT "Molecular characterization of coagulation factor XII-Mie.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095845; BAC23095.1; -.
SQ SEQUENCE 615 AA; 67735 MW; 030508870A0C7EDB CRC64;

Query Match 33.4%; Score 170.5; DB 4; Length 615;
Best Local Similarity 48.5%; Pred. No. 6.2e-12;
Matches 33; Conservative 6; Mismatches 24; Indels 5; Gaps 2;

QY 2 TCYEGNGHFYRGKAS-TDTMGRPCLPWSATVLQOTYHAHRSALQLGLGKHNYCRNPDR 60
DB 216 SCYDGRGLSYRGUARTTLGAPCQPWAS----EATYRNVTAEQARNWGLGSHAFCRNPDR 271
QY 61 RRPWCYV 68
DB 272 DIRPWCYV 279

RESULT 21
QBAXY6
ID QBAXY6 PRELIMINARY; PRT; 947 AA.
AC QBAXY6;

```

```

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
OS Muscle-specific receptor tyrosine kinase MusK.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538710; PubMed=11083926;
RA Ip F.C., Glass D.G., Gies D.R., Cheung J., Lai K.O., Fu A.X.,
RA Yancopoulos G.D., Ip N.Y.;
RT "Cloning and characterization of muscle-specific kinase in chicken.";
RL Mol. Cell. Neurosci. 16:661-673(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Gies D., Glass D.J., Yancopoulos G.D.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY143173; AAN05008.1; -.
KW Receptor; Kinase.
SQ SEQUENCE 947 AA; 105588 MW; E6C3FCC5796CC8BE CRC64;

Query Match 32.1%; Score 163.5; DB 13; Length 947;
Best Local Similarity 39.8%; Pred. No. 6.9e-11;
Matches 38; Conservative 6; Mismatches 27; Indels 27; Gaps 6;

QY 1 KTCYEGNGHFYRGKAS-TDTMGRPCLPWSATVLQOTYHAHR-----SDALQLGLGK 51
DB 463 RTCYSGNGQYVQGWANVTASGIPCQKWS----DQAPHLHRRTPQVPFELSDA----- 510
QY 52 HNYCRNP-DNRRPWCYVQGLKPLV--QECMVHDCAD 86
DB 511 ENYCRNPGENEPWCYTK--DPSVTWYCVSVSPCGD 545

RESULT 22
QB0Y90
ID QB0Y90 PRELIMINARY; PRT; 452 AA.
AC QB0Y90;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE KREMEN.
CN KREMEN.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21167372; PubMed=11267660;
RA Nakamura T., Aoki S., Kitajima K., Funakoshi H., Takahashi T.,
RA Matsumoto K., Nakamura T.;
RT "Molecular cloning and characterization of Kremen, a novel kringle-
RL containing transmembrane protein.";
RL Biochim. Biophys. Acta 1518:63-72(2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR EMBL; AB070851; BAB64294.1; -.
DR InterPro; IPR000859; CUB domain.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR002889; WSC.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00051; kringle; 1.
DR Pfam; PF01822; WSC; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS01180; CUB; 1.

```

Db 157 DTGFWCFT---MDPSVWEYCNLTRCSD 181

RESULT 24

Q9UIR5 QUIRS PRELIMINARY; PRT; 113 AA.

AC QUIRS

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Apolipoprotein(a) (Fragment).

GN APOA.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

RN [1]

XX SEQUENCE FROM N.A.

RP MEDLINE=1181705; PubMed=11285247;

RX RA Ogorelkova M., Kraft H.G., Ehnholm C., Utermann G.;

RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV

RT types 6 to 10 domain affect lip(a) plasma concentrations and have

RT different patterns in Africans and Caucasians.";

RL Hum. Mol. Genet. 10:815-824(2001).

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

DR EMBL; AF158663; AAF03680.1; -.

DR EMBL; AF158662; AAF03680.1; JOINED.

DR HSP; P00747; IEMK.

DR InterPro; IPR000001; Kringle.

DR Pfam; PF00051; Kringle; 1.

DR PRINTS; PR00018; KRINGLE.

DR ProDom; PD000395; Kringle; 1.

DR SMART; SM00130; KR; 1.

DR PROSITE; PS00021; KRINGLE 1; 1.

DR PROSITE; PSS0070; KRINGLE-2; 1.

KW Glycoprotein; Kringle; Lipoprotein.

FT NON TER 1

FT NON TER 113

SQ SEQUENCE 113 AA; 12685 MW; F3D65681D9B5253A CRC64;

Query Match 30.68; Score 156; DB 4; Length 113;

Best Local Similarity 37.8; Pred.No.5.3e-11;

Matches 34; Conservative 12; Mismatches 32; Indels 12; Gaps 4

Qy 1 KTCYEGNGHYRKGKASTDTMGRCPLPWSATV--LQQTYYAHRSDALQLGLGKHYCRNP 58

Db 9 RCQYHGNGQSYRGTFSTVTGRTCSWSSMTPHRHQRTPEYPRDGLTM----NYCRNP 63

Qy 59 DNRRRPCYVQVGLKPLV--QECMVHDCAD 86

Db 64 DADTGPWCFT---MDPSIRWEYCNLTRCSD 90

RESULT 25

O18783 PRELIMINARY; PRT; 806 AA.

AC O18783

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Plasminogen.

OS Macropus eugenii (Tamar wallaby).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.

NCBI_TaxID=9315;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=98004511; PubMed=9342350;

RA Lawn R.M., Schwartz K., Fathly L.;

RT "Convergent evolution of apolipoprotein(a) in primates and hedgehog.";

RT Proc. Natl. Acad. Sci. U.S.A. 94:11992-11997(1997).

CC -1- SIMILARITY: CONTAINS 5 KRINGLE DOMAINS

```
DR EMBL; AF012297; AAB65760.1; -.
DR HSSP; P00747; 5HPG.
DR MEROPS; S01.233; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR003609; Pan app.
DR InterPro; IPR003966; Prothrombin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00051; kringle; 5.
DR Pfam; PF00024; PAN; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PS00722; CHYMOTRYPSIN.
DR PRINTS; PR00018; KRINGLE.
DR PRINTS; PR01505; PROTHROMBIN.
DR ProDom; PD000395; Kringle; 5.
DR SMART; SM00130; KR; 4.
DR SMART; SM00473; PAN_AP; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00021; KRINGLE_1; 5.
DR PROSITE; PS50070; KRINGLE_2; 5.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Kringle; Protease; Serine protease.
SQ SEQUENCE 806 AA; 90981 MW; 95FAA86DC20064D5 CRC64;

Query Match 30.2%; Score 154; DB 6; Length 806;
Best Local Similarity 35.5%; Pred.No. 9e-10;
Matches 33; Conservative 11; Mismatches 25; Indels 24; Gaps 5;

QY 3 CYEENGHYRGKASTDTMGRPCLPWSNATVLQQTVAHRSALQQLGLGKHNYCRNPDNR 54
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 371 CYEKGNGYRGKASTDTMGRPCLPWSNATVLQQTVAHRSALQQLGLGKHNYCRNPDNR 419
QY 55 CRNPDRRRPWCYVOGLKPLV--QECMVHDC 85
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 420 CRNPDRRRPWCYVOGLKPLV--QECMVHDC 449

RESULT 26
ID Q9UIR7 PRELIMINARY; PRT; 113 AA.
AC Q9UIR7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Apolipoprotein(a) (Fragment).
GN APOA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=2118705; PubMed=11285247;
RA Ogorelikova M., Kraft H.G., Ehrholm C., Utermann G.;
RT "Single nucleotide polymorphisms in exons of the apo(a) kringles IV
RT types 6 to 10 domain affect Lp(a) plasma concentrations and have
RT different patterns in Africans and Caucasians.";
RL Hum. Mol. Genet. 10:1815-824 (2001).
CC -!- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.
DR EMBL; AF158659; AAF03678.1; -.
DR EMBL; AF158658; AAF03678.1; JOINED.
DR HSSP; P00747; 2PK4.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00021; KRINGLE_1; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Glycoprotein; Kringle; Lipoprotein.

FT NON_TER 1
FT NON_TER 113
SQ SEQUENCE 113 AA; 12815 MW; 4F80ADF8708548CB CRC64;

Query Match 30.0%; Score 153; DB 4; Length 113;
Best Local Similarity 39.3%; Pred.No. 1.2e-10;
Matches 33; Conservative 8; Mismatches 35; Indels 8; Gaps 3;

QY 3 CYEENGHYRGKASTDTMGRPCLPWSNATVLQQTVAHRSALQQLGLGKHNYCRNPDNR 62
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 11 CYHGDGQSYRGKASTDTMGRPCLPWSNATVLQQTVAHRSALQQLGLGKHNYCRNPDNR 67
QY 63 RPWCYVOGLKPLV--QECMVHDC 84
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 68 RPWCYVOGLKPLV--QECMVHDC 88

RESULT 27
ID Q8WNR1 PRELIMINARY; PRT; 359 AA.
AC Q8WNR1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
SEQUENCE FROM N.A.
RA Pirie-Shepherd S.R., Coffman K.T., Resnick D., Chan R., Kisker O.,
RA Folkman J., Waters D.J.;
RT "Angiostatin is Detectable in the Urine of Dogs with Spontaneous Bone
RT Cancer.";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 4 KRINGLE DOMAINS.
DR EMBL; AY069885; AAL58519.1; -.
DR InterPro; IPR000001; Kringle.
DR InterPro; IPR003966; Prothrombin.
DR Pfam; PF00051; kringle; 4.
DR PRINTS; PR00018; KRINGLE.
DR ProDom; PD000395; PROTHROMBIN.
DR SMART; SM00130; KR; 4.
DR PROSITE; PS00021; KRINGLE_1; 4.
DR PROSITE; PS50070; KRINGLE_2; 4.
KW Glycoprotein; Kringle.
FT NON_TER 1
FT NON_TER 359
SQ SEQUENCE 359 AA; 41172 MW; 776D35F4AB0BDD9E CRC64;

Query Match 29.7%; Score 151.5; DB 6; Length 359;
Best Local Similarity 44.3%; Pred.No. 6.6e-10;
Matches 31; Conservative 5; Mismatches 21; Indels 13; Gaps 3;

QY 3 CYEENGHYRGKASTDTMGRPCLPWSNATVLQQTVAHRSALQQLGLGKHNYCRN 57
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 282 CYHGDGQSYRGKASTDTMGRPCLPWSNATVLQQTVAHRSALQQLGLGKHNYCRN 333
QY 58 PDNRPRPWCY 67
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 334 PDNRPRPWCY 343

RESULT 28
ID Q9R0W3 PRELIMINARY; PRT; 812 AA.
AC Q9R0W3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Plasminogen protein precursor (EC 3.4.21.7).
```


